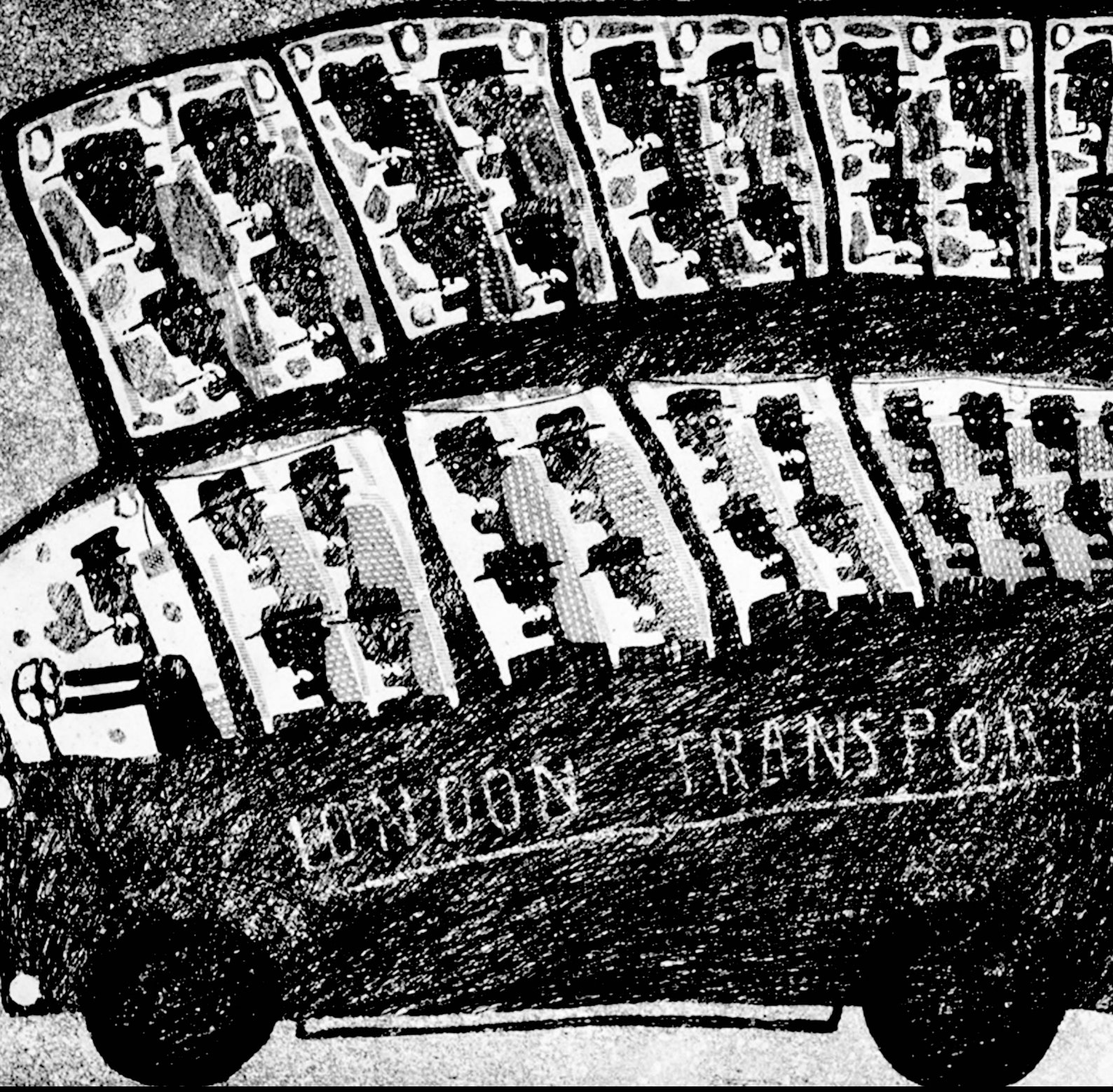



# OFFICIAL CONFERENCE PROCEEDINGS

## The 14th European Conference on Education (ECE2026)

July 09-13, 2026 | London, United Kingdom, and Online

**ISSN: 2188-1162**

Organised by the International Academic Forum (IAFOR) in partnership with the IAFOR Research Centre at the Osaka School of International Public Policy (OSIPP) at The University of Osaka, Japan, and IAFOR's Global Partners.

iafor

"To Open Minds, To Educate Intelligence, To Inform Decisions"

The International Academic Forum provides new perspectives to the thought-leaders and decision-makers of today and tomorrow by offering constructive environments for dialogue and interchange at the intersections of nation, culture, and discipline. Headquartered in Nagoya, Japan, and registered as a Non-Profit Organization (一般社団法人), IAFOR is an independent think tank committed to the deeper understanding of contemporary geo-political transformation, particularly in the Asia Pacific Region.

INTERNATIONAL

INTERCULTURAL

INTERDISCIPLINARY

iafor

# International Academic Board

Professor Anne Boddington, IAFOR, Japan (IAB Chair)
Dr Joseph Haldane, IAFOR & The University of Osaka, Japan
Professor Jun Arima, IAFOR & The University of Tokyo, Japan
Professor Virgil Hawkins, IAFOR Research Centre & The University of Osaka, Japan
Mr Lowell Sheppard, IAFOR & Never Too Late Academy, Japan

Dr Susana Barreto, University of Porto, Portugal
Professor Grant Black, Chuo University, Japan
Dr Evangelia Chrysikou, The Bartlett School of Sustainable Construction, University College London, United Kingdom
Professor Donald E. Hall, Binghamton University, United States
Professor Brendan Howe, Ewha Womans University, South Korea & The Asian Political and International Studies Association (APISA)
Dr James W. McNally, University of Michigan, United States & NACDA Program on Aging

## ECE/ECAH2026 Conference Programme Committee

Dr Rena Alasgarova, Baku Oxford School, Azerbaijan
Professor Emerita Doris Brigitte Ash, University of California Santa Cruz, United States
Professor Anne Boddington, IAFOR, Japan
Dr Evangelia Chrysikou, The Bartlett School of Sustainable Construction, University College London, United Kingdom
Dr Mehmet Demir, University of Birmingham, United Kingdom
Professor Jean-Marc Dewaele, Birkbeck, University of London, United Kingdom
Dr Joseph Haldane, IAFOR & The University of Osaka, Japan
Dr Takoua Jendoubi, University College London, United Kingdom
Dr Rana Khalife, University College London, United Kingdom
Dr Jacqueline Lottin, Higher Colleges of Technology, United Arab Emirates
Dr David Mallows, University College London Institute of Education, United Kingdom
Dr Mauryn Chika Nweke, University College London, United Kingdom
Professor Andrea Révész, University College London Institute of Education, United Kingdom
Dr Ian Scott, University College London, United Kingdom
Dr Marcelo Staricoff, University of Sussex, United Kingdom

## EGen2026 Conference Programme Committee

Dr Jane Biddulph, University College London, United Kingdom
Professor Dimitrios Buhalis, Bournemouth University Business School, United Kingdom
Dr Dorina Cadar, Brighton & Sussex Medical School, United Kingdom
Professor Stefano Capolongo, Polytechnic University of Milan, Italy
Dr Evangelia Chrysikou, The Bartlett School of Sustainable Construction, University College London, United Kingdom (Conference Chair)
Ms Carina Dantas, SHINE 2Europe, Portugal
Professor Eddy Davelaar, Birkbeck, University of London, United Kingdom
Dr Isaiah Durosaiye, School of Architecture, University of Sheffield, United Kingdom
Professor Joseph Falzon, Centre for Research & Innovation, Malta
Professor Ava Fatah, Bartlett School of Architecture, University College London, United Kingdom
Dr Joseph Haldane, IAFOR & The University of Osaka, Japan
Professor Paul Higgs, Faculty of Brain Sciences, University College London, United Kingdom
Dr Fernando Loizides, School of Computer Science & Informatics, Cardiff University, United Kingdom
Dr Elena Petelos, Faculty of Medicine, University of Crete, Greece & Faculty of Health, Medicine & Life Sciences, Maastricht University, Netherlands
Ms Eleftheria Savvopoulou, SynThesis Architects, Greece
Professor Anastasios Tellios, School of Architecture, Aristotle University of Thessaloniki, Greece
Dr Greg Williams, University of Manchester, United Kingdom

The European Conference on Education 2026

Official Conference Proceedings

ISSN: 2188-1162

Executive Editor:
Joseph Haldane, PhD

IAFOR Publications Manager:
Nick Potts

IAFOR Publications Coordinator:
Mark Kenneth S. Camiling

Editorial Assistants:
Sachiko Sakamoto
Suzue Abe
Yuika Sakaguchi
Arisa Egami

Date Published: September 23, 2026

Published by The International Academic Forum on the IAFOR Research Archive
URL: https://papers.iafor.org/

© The International Academic Forum 2026
The International Academic Forum (IAFOR)
Sakae 1-16-26 2F
Naka Ward, Nagoya, Aichi
Japan 460-0008
iafor.org

## Table of Contents

**Role-Playing Games as a Psychometric Tool for the Analysis of Personality Profiles and Academic Performance in Basic Education**
Waldir Siqueira Moura
Edgar Delbem
Angélica Fonseca da Silva Dias
Juliana Baptista dos Santos França pp. 1 - 14

**Augmented Reflection in Practicum: AI as a Dialogic Partner in Teacher Education**
Anna Jo Perry pp. 15 - 23

**Maker Culture and Sustainability in Basic Education**
Lígia Maura Delbem
Waldir Siqueira Moura
Edgar Delbem
Célia Regina Tomachuk dos Santos Catuogno pp. 25 - 37

**"AI Talent for All": Sociotechnical Imaginaries in Korea's 2025 AI Education and Regulation Strategy**
Seri Yoon pp. 39 - 52

**Surpass Chatbots and Human Teachers? Explore the Effectiveness of Conversational AI in Language Learning**
Yitian Sun
Shimeng Wu
Cong Qi pp. 53 - 70

**Children as Guides in Storytelling: Exploring Imagination and Expression in Child–AI Collaborative Narrative Play**
Ting Xu
Xian Xu pp. 71 - 83

**A Competency-Oriented System of Chemistry Tasks for Grade 7: Scientific and Methodological Foundations and Empirical Validation of Effectiveness**
Indira Aimbetova
Dinara Berdi
Akylbek Meirbekov pp. 85 - 101

**Pedagogical Technology Integration in ICT-Rich and ICT-Poor Community Learning Centres in South Africa**
Khethokwakhe Goodman Chonco
Khanyisile Mbatha pp. 103 - 114

**The AI Academic Integrity Toolkit: Designing Ethical and Reflective Assessment Practices in the Age of AI**
Rahme Sadikoglu pp. 115 - 127

**AI Integration in Pedagogical Practices: A Quantitative Analysis of the Competence of College Teachers**
Joan P. Guinto
Shayne Klarisse E. Dinamling
Ibrahim F. Hanbal
Don Harl C. Malabanan
Dionisio R. Tandingan Jr. pp. 129 - 137

**Time Series Forecasting Using Context Rich Restaurant Sales Data in Data Science Education**
Boyang Zhang pp. 139 - 148

**Critical Reflection and Inclusive Education: A Semester-Long Implementation of Brookfield's Four Lenses in Teacher Education**
Natassa Raikou
Labrina Gioti pp. 149 - 161

**A Multi-model Deliberative Generative AI Framework: Enhancing Critical Thinking and Content Veracity in Higher Education**
Shyh-Jian Tan
Y. C. K. Chen pp. 163 - 169

**Building an AI Teaching Assistant With RAG: Implementation and Performance Scoring Framework**
Alexandre Moudjeb
Fahmi Bellalouna pp. 171 - 187

**Instructional Leadership in Underperforming Schools: Evaluating the Intervention Strategies of School Management Teams and Subject Advisors**
Sandile Xaba
Richard Moloele
Maseeeng Papashane pp. 189 - 203

**Development of Inclusive Education in Nigeria**
Ugonne Ezugwu pp. 205 - 220

**Leading Differently: Experiences of Neurodivergent Education Leaders and Their Professional Journeys**
Daniel Beck pp. 221 - 245

**"A Very Good Girl": Experiencing Primary School After Intrafamilial Sexual Abuse From an Adult Male**
Sheona Goodyear pp. 247 - 257

**Beyond the Classroom: Industry-Mentored Experiential Workshops Transform Interior Design Education**
Jeannine Vail
Kathleen Branning pp. 259 - 271

**Towards Immersive Training for Inclusive Education: Co-designing a Virtual Reality Application for Educators**
Malak Kanaan
Sylvain Fleury
Vivien Walter
Karine Galmiche
Simon Richir
Nour El Mawas pp. 273 - 285

**A Tale of Two Assessments: Navigating Generative AI Integration in Authentic Assessment Design**
Ramalakshmi Vaidhiyanathan pp. 287 - 292

**Does More Coursework Mean Better Preparation? Examining Course Availability and Opportunities to Learn Inclusive Practices in Teacher Education**
Melike Yozgatli Gulum pp. 293 - 306

**An Exploration of the Use of Musical Activities to Promote Foundational Literacy Competencies for Children Aged 3 to 5 Years**
Maria Kay pp. 307 - 325

**From "Curriculum as Product" to "Curriculum as Process": An Autoethnographic Critical Review of English Language Curriculum Development in Hong Kong**
Shuk Ling Cheng
Wenyun Jia pp. 327 - 333

**Spaces That Follow Individual Preference Can Improve Learning Performance**
Walaa Aljehani
Ana Rute Costa
Michelle To pp. 335 - 347

**Designing and Facilitating External Partner-Based Pedagogy in Higher Education: Implications for Authentic Learning, Student Engagement, and Perceived Relevance**
Rachel Masson
Ana Da Silva Costa
Jed Lennox pp. 349 - 357

**The Early Childhood Education Crisis Across Socioeconomic Contexts: A Comparative Qualitative Study From Israel**
Amit Rottman
Yoad Eliaz
Lior Gendelman pp. 359 - 371

**Visualizing the Invisible: Co-creating an AI-Assisted Tactile Picture Book With a Congenitally Blind Student**
Chu-Yu Cheng
Kuan-Ju Ho pp. 373 - 384

**Reflective Practice in Pre-professional Arts Education: A Theoretical Examination Through Teacher Education**
Misun Kong
Minsung Kwon
Jonghun Kim pp. 385 - 397

**Reconstructing Lexical Networks of Japanese EFL Learners: An Application of Technique Feature Analysis to Already-Known Words**
Noriko Aotani
Shin'ya Takahashi pp. 399 - 407

**Strengthening Self-Regulation and Executive Functions Through a Value-Based Challenge-Board Framework: Educational Implications for Contemporary Schools**
Stephan Drescher pp. 409 - 421

**The Price of Progress? Assessing the Impact of Artificial Intelligence on Systems Engineering Education**
Kieran Holling
Steve Barker pp. 423 - 433

**Leveraging Human Intelligence: Assessment for Learning Through Peer Feedback and Self-Reflection**
Negar Riazifar
Edwina Jones
Sam Grierson
Kerry Dobbins pp. 435 - 447

**The Role of Academic Capital in Reducing Social Asymmetries in Brazilian Universities**
Wescley Xavier
José Roberto Carvalho Junior
Ykaru Wagner
Ronald Gonçalves pp. 449 - 461

**Designing Game-Based Online Courseware for Workplace Chinese Communication in a University Context**
Hon Fong Poon pp. 463 - 468

**Individual Conceptions of Teaching of Prospective Pre-primary Teachers in Relation to the Integration of Educational Content**
Ivana Rochovská pp. 469 - 481

**Beyond GenAI Adoption: New Horizons for the Ethos of Teaching in Higher Education**
Adma Gama-Krummel pp. 483 - 497

**Implementation of Integrated Curriculum in the Opinions of Primary Education Teachers**
Monika Brozmanová
Ivana Rochovská pp. 499 - 508

**Designing for Disciplinary Diversity: A Comparative Analysis of Preferences in Informal Learning Environments**
Siriwan Rujibhong pp. 509 - 523

**AI Literacy, Collaboration, and Future STEM Interests: An Informal Seven-Week Dual-Enrollment Presidential AI Challenge**
Elodie Billionniere
Lawrence Meyer, Jr. pp. 525 - 542

**The Pedagogical Effectiveness of Integrating Gamification and AI Tools in Teaching Physics**
Indira Usembayeva
Gulzhan Niyazova pp. 543 - 557

**Understanding Sustainability in the Context of Anchoring and Objectification: Insights From College Students**
Özden Şengül pp. 559 - 565

**How Much Does Financial Support Matter? Changes in Study Abroad Participation and Destination Choice at a Japanese University**
Daniel James pp. 567 - 578

**Does AI Ideation Feel Collaborative? Collaboration Potential as a Structurally Distinct Evaluation Dimension in an LLM-Based Ideation Tool**
Stefania Zourlidou
Kamyab Farokhi
Frank Hopfgartner pp. 579 - 585

**Design Conditions for Educator Adoption of LLM-Based Ethics Scaffolding in Higher Education**
Stefania Zourlidou pp. 587 - 596

**3 Level Game for the First-Year Students in Computer Science**
Daniela Tsaneva pp. 597 - 604

**Evaluating Simulation-Based Learning in Management Education: Evidence From Student Perceptions**
Salman Ali pp. 605 - 611

**Industry-Integrated Curriculum Design for Employability in Marketing Education: The BRIDGE Process Framework**
Ying Wu
Kirsty Beaumont
Ruth O'Riordan pp. 613 - 636

**Designing AI-Supported Assessment in Marketing Education: Specified Latitude and the CRAFT Framework**
Ying Wu
Andrew Ross
Malcolm Stewart
Fahima Alam pp. 637 - 658

**Beyond Exposure: Emotional Attachment in Heritage Language Retention Among Bilingual Families**
Venera Kildyusheva
Anna Volkova pp. 659 - 672

**The Experience of Mexican Female Students in the COIL Mexico-Colombia Project: Insights Into Intercultural Online Collaboration and Educational Exchange**
Juana Beatriz Ortiz-Cienega
Monica Mondelo-Villasenor
Bertha Mondelo-Villasenor pp. 673 - 678

**From Majlis to Manuscript: Leveraging Emirati Oral Traditions to Scaffold Phase 1 Literacy in the PYP**
Nguyilan Shaaji
Alueshima Jennifer Shaapera pp. 679 - 690

**Total Quality Management in Action: What Headmasters of High School Believe**
Vasileios Tsiantos
Argiris Fyntanakis pp. 691 - 702

**Overview of Industrial Engineering Courses in Brazil Under the Focus of Entrepreneurship**
Adriana de Paula Lacerda Santos
Amanda Santos
Bruno Lacerda Santos pp. 703 - 709

**“It’s a Private and Shameful Thing”: Young People With Intellectual Disabilities Making Sense of Unwanted Sexual Messages on Social Media**
Cansu Bolat Dinc
Fatma Nur Akcin pp. 711 - 727

**Bridging the Distance: Leveraging Communication Apps to Foster Collaboration and Reduce Isolation in Online Graduate Education**
Gina Solano
Jennifer Miller pp. 729 - 744

**The Homework as a Space for Dialogue in Curriculum Design**
Barbara Bocchi pp. 745 - 758

**Coherence as a Coordination Problem: The Coherence Navigation Framework for Multi-agency Support in Neurodivergent Children's Education**
Azibaolanari Okungbowa pp. 759 - 768

**Increasing Intraverbal Response Order Variability in a Child With Autism: Effects of Lag Schedule Combined With Differential Reinforcement**
Masashi Kawamura pp. 769 - 779

**Supporting Pasifika Learners in New Zealand Tertiary Education: Strengths-Based Pastoral Care Practices to Enhance Pasifika Learner Wellbeing and Educational Outcomes**
Krishan Mani
Anamua Lole pp. 781 - 788

**Designing Career Readiness Through Industry-Embedded Capstone Pedagogy: A Curriculum Design Framework for Professional Identity Formation**
Barbara Trippeer pp. 789 - 805

**Non-cognitive Factors of Mathematics Achievement Among Higher Education Students in Bahrain**
James Ray Mendaje pp. 807 - 830

**Migration Intentions Among Mathematics Educators in Higher Education Institutions (HEIs) in Batangas, Philippines**
Lea Catapang pp. 831 - 847

**Technology-Mediated English Grammar Learning by Chinese Learners: A Case Study**
Chi Wai Lee pp. 849 - 857

**Playing Heritage: Loose Parts and Traditional Emirati Play as Pathways to Culture and Identity in Early Childhood**
Alueshima Jennifer Shaapera
Nguyilan Shaaji pp. 859 - 868

**Bridging Policy and Practice in India: Implementing IB Inquiry-Based Pedagogy in a Tier-3 CBSE School**
Ramya Damarla pp. 869 - 883

**Type One Activations in Suggestopedia: Scaffolding, Suggestion, and the Activation of Hidden Reserves**
Christina Johnstone pp. 885 - 897

**The DIAL Framework for AI-Supported Decision Making in Undergraduate Business Programs**
Eduard Merc pp. 899 - 905

**Development and Evaluation of a Gamified Digital Platform for Science Education in Secondary Schools**
Nurseit Zhunissov
Anuarbek Amanov
Arypzhan Aben pp. 907 - 918

**A Comparative Study on Oral Communication Strategy Use Among Vietnamese Non-English Majors**
Huyen Nhi Vo
Huu Tai Nguyen pp. 919 - 936

**Safety Experience and Student Well-being at School: The Moderating Role of School Characteristics Among Indonesian Students**
Septian Cahyo
Siti Raudlah
Januar Pribadi pp. 937 - 949

**Evaluating the Effectiveness of Gamification Technology in Teaching Science Subjects**
Anuarbek Amanov
Nurseit Zhunissov
Arypzhan Aben pp. 951 - 959

**English Production Learning Cycle Using Generative AI in Image Description Tasks: Focusing on Lowering Psychological Barriers and Promoting Autonomous Noticing**
Harumi Kashiwagi
Min Kang pp. 961 - 974

**Exploring the Use of a Culturally Responsive Pedagogy in Practice: Perception and Knowledge Performance With the CARE² Model**
Karen K. Fujii
Natalie Perez pp. 975 - 988

**Process-Oriented Collaborative Essays: Empowering Undergraduate Students Through Assessment Redesign in the Age of Generative AI**
Fahima Alam
Jonathan Jackson pp. 989 - 1002

**Artful Early Years: Children's Gendered Awareness and Expression**
Michael Brain Martins pp. 1003 - 1027

# Role-Playing Games as a Psychometric Tool for the Analysis of Personality Profiles and Academic Performance in Basic Education

Waldir Siqueira Moura, Federal University of Rio de Janeiro, Brazil
Edgar Delbem, Federal University of Rio de Janeiro, Brazil
Angélica Fonseca da Silva Dias, Federal University of Rio de Janeiro, Brazil
Juliana Baptista dos Santos França, Federal University of Rio de Janeiro, Brazil

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This study investigates whether structured decisions made in a role-playing game can generate psychometrically interpretable student profiles associated with academic performance in basic education. The proposed Psychogenesis-RPG is an analogue game-based assessment comprising 40 character-customisation items distributed across eight behavioural dimensions. An observational and correlational design was employed with students from the final years of basic education, while an age-eligible subsample completed the 16 Personality Factor Questionnaire (16PF) to examine convergent evidence. The analytical framework combines ordinal scoring, internal consistency, rank-order correlations, and profile comparison. Results indicate that game-derived behavioural profiles provide meaningful educational information and complement conventional academic indicators without replacing standardised psychological assessment. The study contributes to educational psychometrics and Learning Analytics by proposing a practical framework for formative student profiling through structured role-playing decision-making.

*Keywords*: role-playing games, game-based assessment, educational psychometrics, personality profiles, academic performance, student profiling, basic education, learning analytics

iafor
The International Academic Forum
www.iafor.org

## Introduction

Academic performance is one of the most visible indicators used to evaluate student progress in basic education. Grades, test scores, attendance records, and completion rates provide important evidence about learning outcomes, but they reveal comparatively little about the behavioural processes through which those outcomes are produced. Two students with similar grades may differ substantially in planning, persistence, autonomy, willingness to collaborate, tolerance for uncertainty, or capacity to adapt when a task becomes difficult. Conversely, students with similar behavioural tendencies may attain different outcomes because of contextual, instructional, or socioeconomic conditions. Academic performance should therefore be interpreted not as an isolated expression of ability, but as the result of interactions among cognitive, motivational, behavioural, social, and environmental factors.

This multidimensional perspective has increased interest in student profiling within Educational Data Mining and Learning Analytics. These fields seek to transform traces of learning activity into interpretable evidence capable of supporting feedback, monitoring, instructional adaptation, and early intervention (Baker & Inventado, 2014; Siemens & Long, 2011). In digital learning environments, profiling commonly relies on interaction logs, task-completion sequences, response times, navigation patterns, and participation records. Such evidence can reveal regularities that are not immediately visible in final grades. Nevertheless, the implementation of data-intensive educational analytics remains uneven in basic education, particularly in institutions with limited technological infrastructure or restricted access to specialised assessment systems.

The practical value of student profiling also depends on how profiles are conceptualised. When behavioural indicators are treated as fixed categories, they may reinforce deterministic interpretations, reduce complex students to simplified labels, or encourage educational decisions based on unstable evidence. A more appropriate formative approach treats profiles as provisional representations of observable tendencies within a specific context. From this perspective, profiling is not intended to establish who a student permanently is. Instead, it provides hypotheses about how the student tends to approach decisions, interact with others, organise tasks, or respond to uncertainty under the conditions in which the evidence was collected.

Game-based assessment offers a promising means of producing such evidence. Games require participants to make decisions under rules, allocate resources, select strategies, respond to consequences, and negotiate relationships between immediate and delayed objectives. Role-playing games are particularly relevant because they situate decisions within a narrative and invite players to express preferences through character construction, group roles, strategic choices, and responses to hypothetical events. These features create a structured but contextually meaningful environment in which behavioural tendencies can be elicited without reproducing the format of a conventional personality questionnaire.

The educational use of games has traditionally emphasised motivation, engagement, content learning, and collaborative problem solving. Research on gamification and game-based learning indicates that appropriately designed game elements can influence participation and motivation, although their effects depend on instructional context, learner characteristics, and alignment between mechanics and pedagogical objectives (Deterding et al., 2011; Hamari et al., 2014; Sailer et al., 2017). Less attention has been given to the use of games as measurement environments capable of producing psychometrically interpretable indicators. In such

environments, choices are not merely part of gameplay; they constitute observable responses from which inferences about tendencies or constructs may be developed.

This measurement perspective requires caution. A decision made in a role-playing task cannot automatically be interpreted as evidence of a stable personality characteristic. The relationship between game behaviour and psychological constructs depends on the design of the task, the clarity of the response alternatives, the consistency of item functioning, and the theoretical basis connecting observable choices to the intended dimension. Game-based assessment must therefore be supported by explicit scoring rules, standardised administration, reliability evidence, and empirical relationships with external variables.

Psychometric instruments such as the 16 Personality Factor Questionnaire provide established procedures for assessing personality-related constructs. Their standardisation, normative frameworks, and professional administration requirements support controlled psychological interpretation. At the same time, these characteristics may limit their routine use as formative educational tools, particularly when schools seek rapid, accessible, and context-sensitive evidence for pedagogical planning. Game-based instruments should not be viewed as substitutes for standardised personality inventories. Their potential lies instead in generating complementary indicators under ordinary educational conditions and in identifying patterns that may warrant closer pedagogical observation.

The present study examines this possibility through Psychogenesis-RPG, an analogue role-playing assessment based on structured character-customisation decisions. The instrument comprises 40 forced-choice items distributed across eight dimensions: Avatar Identity and Aesthetics, Preferred Role in Team/Narrative, Strategic Orientation, Tolerance for Risk and Uncertainty, Self-Control and Planning, Sociability and Interaction, Exploration and Curiosity, and Empathy and Other-Focus. Each item presents four ordered alternatives representing increasing levels of the tendency associated with its target dimension. Responses are aggregated into dimension-level scores intended to describe relative differences among participants rather than establish clinical or diagnostic categories.

Unlike research centred exclusively on the psychometric validation of a new instrument, this study focuses on the educational meaning of the profiles generated through RPG decisions. Its central concern is whether variation in these profiles is associated with students' academic performance and whether particular dimensions provide information that is not apparent from grades alone. For example, Self-Control and Planning may be related to the organisation and completion of academic tasks, while Exploration and Curiosity may be associated with engagement in open-ended activities. Sociability and Interaction may have particular relevance in collaborative learning contexts, whereas Tolerance for Risk and Uncertainty may influence how students respond to unfamiliar or complex tasks.

These relationships should not be assumed to be uniform or linear. A high score in a particular dimension is not necessarily advantageous in every educational context. Greater tolerance for uncertainty may support experimentation and problem solving, but may also be associated with impulsive decisions when planning is limited. Strong self-control may support consistency and task completion, yet excessive rigidity may reduce adaptability in activities requiring exploration. Similarly, sociability may facilitate collaborative work without necessarily predicting performance in individual assessments. The value of a dimensional profile therefore lies in the configuration of tendencies and their interaction with instructional demands, rather than in the presumption that higher scores always represent superior educational functioning.

The analysis of academic performance also demands methodological restraint. Grades are influenced by curriculum, teacher expectations, assessment formats, attendance, prior knowledge, learning opportunities, and institutional practices. Associations between RPG-derived profiles and academic results cannot be interpreted as causal effects or direct measures of intelligence. Instead, they provide evidence about whether students occupying different positions in behavioural dimensions also tend to differ in selected educational indicators. Such relationships may support the formative interpretation of the instrument, but they do not establish that personality profiles determine academic achievement.

To examine the psychometric plausibility of the game-derived indicators, the study combines internal and external forms of evidence. The complete school sample supports the examination of response distributions, dimensional variability, reliability, and associations with academic indicators. An age-eligible subsample is additionally compared with theoretically corresponding dimensions of the 16PF. Because the instruments employ different measurement systems, the comparison is based on rank-order relationships rather than direct score equivalence. The original validation sample comprises 44 students aged between 14 and 17 years, while direct comparison with the 16PF is restricted to the 13 participants meeting its recommended age conditions.

By integrating psychometric analysis and academic evidence, this study addresses a gap between game-based assessment and school-based decision support. Existing work frequently examines whether games improve learning or engagement. The present investigation asks a different question: whether decisions produced within a role-playing structure can function as interpretable data about how students approach educational situations. This distinction positions the RPG not only as an instructional resource but also as a structured environment for eliciting behavioural evidence.

## Research Objective

The general objective of this study is to investigate whether structured decisions made in a role-playing game can generate reliable and educationally interpretable personality-related profiles associated with academic performance in basic education.

More specifically, the study examines:

- the internal functioning and reliability of the eight Psychogenesis-RPG dimensions;
- the distribution of game-derived profiles across students and school grades;
- the relationships between dimensional scores and academic-performance indicators;
- the convergence between selected Psychogenesis-RPG dimensions and theoretically corresponding 16PF factors;
- the potential and limits of using RPG-derived profiles for formative educational decision-making.

## Research Questions

The study is guided by the following research questions:

- RQ1. To what extent do structured role-playing decisions generate internally coherent and differentiated student profiles?
- RQ2. Which Psychogenesis-RPG dimensions are associated with academic performance in basic education?

- RQ3. Do students with different global formative classifications exhibit distinct academic-performance patterns?
- RQ4. Do game-derived dimensions present rank-order convergence with theoretically corresponding factors of the 16PF?
- RQ5. How can RPG-derived profiles complement conventional academic indicators without producing deterministic or diagnostic classifications?

## Contributions

This study makes four principal contributions:

1. First, it reframes the role-playing game as a psychometric data-generation environment, rather than treating it exclusively as a motivational or instructional strategy. The decisions embedded in character construction are operationalised as observable indicators linked to explicit behavioural dimensions.
2. Second, it connects game-derived profiling to academic-performance analysis. This connection enables examination of whether dimensions such as planning, exploration, sociability, empathy, and tolerance for uncertainty provide complementary information about student functioning.
3. Third, the study proposes an analytical framework that integrates ordinal scoring, dimensional reliability, external convergence, academic indicators, and profile comparison. The framework is designed for pilot and school-based research contexts in which sample sizes, infrastructure, and access to standardised psychological instruments may be restricted.
4. This study establishes clear ethical and interpretive boundaries. Psychogenesis-RPG profiles are treated as context-dependent formative indicators. They are not intended to define fixed personality types, establish diagnoses, determine educational placement, or replace professional psychological assessment. Their purpose is to enrich pedagogical understanding by combining structured game decisions with other sources of educational evidence.

## Related Work and Theoretical Foundations

Research on digital games in education has traditionally focused on their capacity to improve motivation, engagement, and learning outcomes. Over the last two decades, numerous studies have demonstrated that appropriately designed game environments can increase participation, persistence, and enjoyment while supporting the acquisition of disciplinary knowledge. These effects, however, are neither universal nor independent of instructional context. Their magnitude depends on the alignment between game mechanics, learning objectives, student characteristics, and pedagogical implementation (Deterding et al., 2011; Hamari et al., 2014; Sailer et al., 2017).

As the field matured, researchers began recognising that educational games produce considerably more information than final scores alone. Every decision made during gameplay constitutes a behavioural event capable of revealing aspects of planning, exploration, collaboration, persistence, adaptation, and problem-solving. Consequently, games gradually evolved from being viewed exclusively as instructional interventions to becoming environments capable of generating educational evidence.

This transition gave rise to the concept of Game-Based Assessment, in which gameplay is treated as a structured source of observable behavioural indicators rather than merely a vehicle

for content delivery. Within this perspective, player choices become assessment data, enabling researchers to investigate constructs that are difficult to observe through conventional examinations or self-report questionnaires. Unlike traditional tests, game-based assessments capture behaviour while decisions are being produced, thereby reducing dependence on retrospective self-evaluation and increasing ecological validity.

The emergence of stealth assessment further expanded this perspective. Instead of interrupting gameplay to administer external questionnaires, assessment is embedded directly within the game mechanics, allowing behavioural evidence to accumulate naturally throughout interaction (Shute, 2011). Although many stealth-assessment systems rely on complex computational models and digital log data, their underlying principle is broader: meaningful behavioural information can be inferred from structured decision-making when supported by an explicit measurement framework.

Despite these advances, relatively few studies have investigated analogue games as psychometric environments. Most published work continues to focus on digital platforms because they facilitate automatic data collection and large-scale analytics. Nevertheless, analogue games possess important methodological advantages in school settings. They require minimal technological infrastructure, are easily implemented in classrooms with limited resources, permit direct observation of participant behaviour, and reduce technological barriers that frequently affect digital educational interventions. These characteristics make analogue role-playing games particularly attractive for basic education, where feasibility often determines whether an assessment instrument can be adopted routinely.

## Role-Playing Games as Behavioural Decision Environments

Role-playing games occupy a distinctive position within game-based education because they organise player behaviour around narrative decision-making rather than competition alone. Instead of rewarding isolated correct answers, RPGs require participants to construct identities, assume social roles, negotiate objectives, manage uncertainty, and make successive decisions whose consequences emerge throughout the narrative.

From an educational perspective, this structure creates repeated opportunities for observing behavioural tendencies under controlled conditions. Decisions concerning leadership, cooperation, planning, exploration, empathy, or risk-taking are not requested through direct questions but emerge from contextualised situations requiring participants to select among multiple plausible alternatives. Consequently, RPGs may reveal patterns that remain less evident when students simply describe themselves through conventional personality inventories.

The educational literature has predominantly examined RPGs as instructional strategies. Previous investigations have reported improvements in motivation, creativity, collaboration, historical understanding, language development, and problem-solving. These studies generally evaluate learning gains or participant engagement rather than the psychometric properties of decisions generated during gameplay. The possibility that RPG decisions themselves may constitute measurable behavioural indicators has received considerably less attention.

This distinction is fundamental. When an RPG is designed for learning, decisions primarily serve the progression of the narrative. When designed for assessment, the narrative becomes a structured context through which theoretically relevant behaviours can be elicited

systematically. Under these conditions, each decision contributes to a predefined measurement model, allowing individual responses to be aggregated into interpretable dimensions representing broader behavioural tendencies.

Psychogenesis-RPG adopts this second perspective. Character customisation is not merely an aesthetic component of gameplay but the central assessment mechanism through which students express preferences regarding identity, planning, collaboration, uncertainty, exploration, and interpersonal interaction. Each response therefore functions simultaneously as a narrative choice and as an observable behavioural indicator. The resulting profile is interpreted not as a fictional description of the created character but as evidence concerning the student's preferred decision patterns within the structured environment provided by the assessment.

## Personality-Related Tendencies and Academic Performance

Academic achievement is influenced by a wide range of interacting variables that extend beyond cognitive ability alone. Research in educational psychology consistently demonstrates that planning, self-regulation, persistence, cooperation, motivation, curiosity, emotional regulation, and adaptive decision-making contribute substantially to students' educational trajectories. These characteristics influence how learners organise their study routines, respond to challenges, collaborate with peers, and maintain engagement throughout the learning process.

Personality research provides one framework for understanding these behavioural differences. Trait models describe relatively stable patterns through which individuals perceive situations, regulate behaviour, and interact with their environment. Although no single personality dimension determines educational success, combinations of behavioural tendencies frequently influence how students approach academic tasks.

For example, students exhibiting stronger planning and self-regulation often demonstrate greater consistency in homework completion and long-term projects. Curiosity and exploratory behaviour may facilitate inquiry-based learning and creative problem-solving. Empathy and sociability can influence participation in collaborative activities, whereas tolerance for uncertainty may become particularly relevant during open-ended investigations requiring experimentation and adaptation.

These relationships should not be interpreted deterministically. Educational performance emerges through continuous interactions among individual characteristics, teaching practices, curriculum design, family context, socioeconomic conditions, and institutional opportunities. Behavioural tendencies therefore represent one component of a considerably broader educational system rather than isolated predictors of achievement.

Accordingly, the present study adopts a probabilistic rather than deterministic interpretation of personality-related profiles. The objective is not to predict grades directly from RPG decisions but to examine whether systematic associations emerge between behavioural dimensions and educational indicators. Such associations may support pedagogical reflection while recognising that educational outcomes remain context-dependent and subject to multiple interacting influences.

## Educational Psychometrics Beyond Conventional Questionnaires

Psychometric assessment in educational contexts has traditionally relied on questionnaires, rating scales, structured inventories, and standardised psychological tests. These instruments offer important advantages, including established theoretical foundations, normative interpretation, and well-developed statistical methodologies. However, they also depend heavily on self-report, verbal comprehension, and respondents' capacity to evaluate their own behaviour accurately.

Game-based assessment introduces a complementary measurement paradigm. Instead of asking participants how they usually behave, it observes the decisions they make while interacting with structured situations. The resulting evidence is behavioural rather than declarative. This distinction does not eliminate measurement challenges, but it provides an alternative source of information that may complement conventional psychometric procedures.

The use of behavioural evidence also aligns with contemporary conceptions of validity, which emphasise that score interpretation should be supported by multiple sources of evidence rather than by a single statistical coefficient. Response processes, internal score structure, relationships with external variables, and theoretical coherence collectively contribute to the validity argument. Consequently, game-derived measures should not be evaluated solely according to whether they reproduce questionnaire scores. Their value depends on whether the behavioural indicators they generate correspond coherently to theoretically meaningful constructs and contribute useful information for the intended educational purpose.

Within this framework, Psychogenesis-RPG is not proposed as an alternative personality inventory. Rather, it functions as a structured behavioural assessment whose educational utility depends on three complementary conditions: internally coherent dimensions, theoretically interpretable behavioural indicators, and meaningful relationships with external educational variables. These variables include both established psychometric measures and academic-performance indicators, allowing the instrument to be evaluated simultaneously from psychological and educational perspectives.

## Research Gap and Conceptual Framework

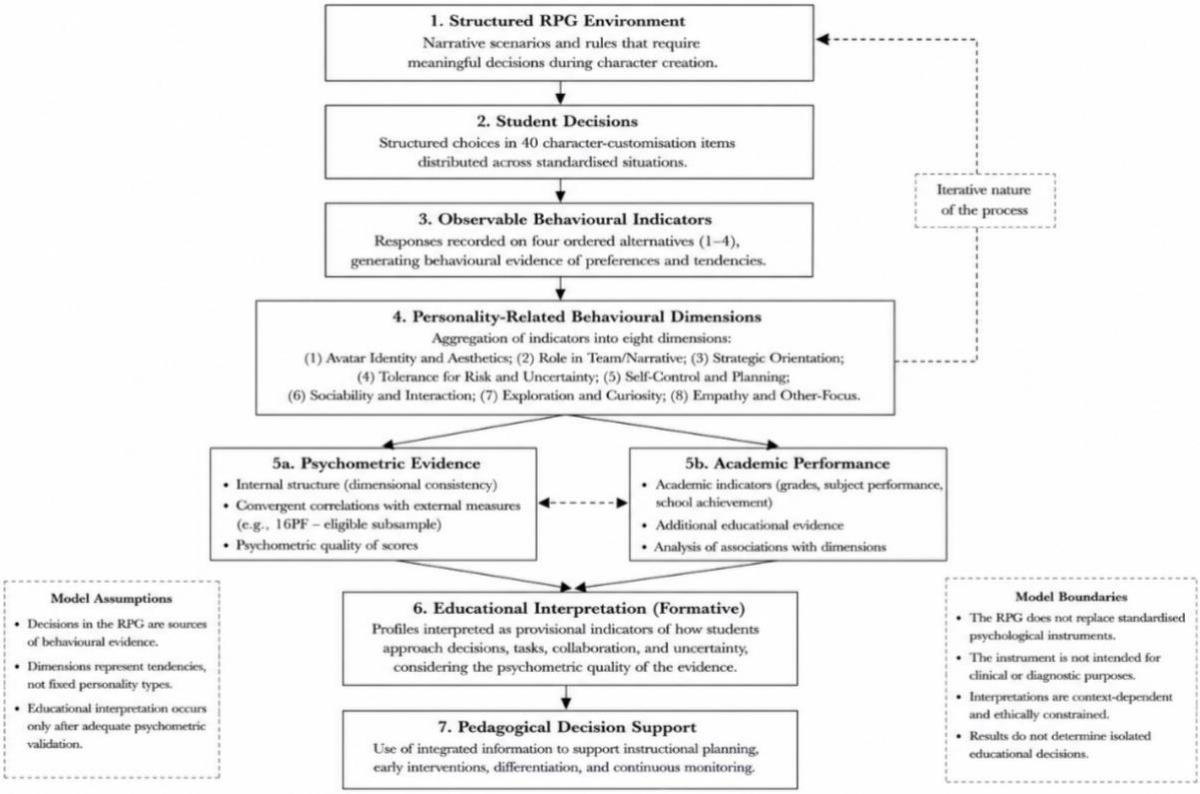
The literature reviewed above reveals an important gap. Existing research demonstrates that games can enhance motivation, engagement, and learning. Educational psychometrics demonstrates how behavioural constructs may be measured reliably. Learning Analytics shows the value of behavioural data for educational decision-making. Nevertheless, these research traditions have developed largely independently.

Few studies investigate whether structured role-playing decisions can simultaneously satisfy psychometric requirements and generate educationally meaningful profiles associated with academic performance. Even fewer examine analogue RPGs designed specifically as behavioural assessment instruments suitable for routine classroom implementation.

The present study addresses this gap by integrating these three perspectives into a single analytical framework. Rather than treating games as instructional interventions alone or psychometrics as questionnaire-based assessment alone, the proposed framework conceptualises the role-playing environment as a controlled source of behavioural evidence. Student decisions are transformed into dimensional indicators through predefined scoring

procedures, examined psychometrically through reliability and convergent analyses, and interpreted educationally through their relationships with academic performance.

**Figure 1**
*Conceptual Model From the RPG Environment to Educational Interpretation*




The figure 1 summarises the conceptual model adopted throughout this investigation. The model proposes that structured RPG decisions generate observable behavioural indicators, which are aggregated into personality-related dimensions. These dimensions constitute the primary analytical unit and are subsequently examined in relation to academic performance and external psychometric evidence. Educational interpretation is therefore situated at the final stage of the process and remains explicitly bounded by the psychometric quality of the generated scores.

## Research Design, Analytical Framework and Findings

This study adopted an observational, correlational, and cross-sectional research design to investigate whether behavioural decisions produced during a structured role-playing game are associated with measurable differences in students' academic performance. Unlike intervention studies that evaluate learning gains following gameplay, the present investigation conceptualises the role-playing environment as a behavioural assessment context in which each decision constitutes observable evidence of individual tendencies relevant to educational functioning.

The research was conducted in a private educational institution offering both the final years of Basic Education and Upper Secondary Education. Participation was voluntary and authorised by the school administration, students, and their legal guardians whenever required. All procedures complied with ethical principles governing educational research involving human

participants. The assessment formed part of a broader investigation into game-based psychometric profiling and educational decision support.

The complete sample consisted of 44 students aged between 14 and 17 years, representing different school years within Basic Education. Because one component of the study involved comparison with the Sixteen Personality Factor Questionnaire (16PF), only participants meeting the recommended age requirements for that instrument composed the convergent-validation subsample (n = 13). Consequently, analyses involving academic performance and behavioural profiling used the entire sample, whereas direct psychometric convergence with the 16PF was restricted to the eligible subgroup.

This distinction reflects the different purposes of the analyses. School-wide profiling sought to examine relationships between game-derived behavioural tendencies and educational indicators, while convergent analysis investigated whether the dimensions extracted from the RPG demonstrated theoretically coherent relationships with an established personality assessment.

## Psychogenesis-RPG Assessment Framework

The assessment environment was based on Psychogenesis-RPG, an analogue role-playing instrument specifically designed to elicit behavioural preferences through structured character customisation rather than through explicit self-report.

Participants completed 40 forced-choice items, organised into eight behavioural dimensions, each represented by five independent decision scenarios:

- Avatar Identity and Aesthetics;
- Preferred Role in Team and Narrative;
- Strategic Orientation;
- Tolerance for Risk and Uncertainty;
- Self-Control and Planning;
- Sociability and Interaction;
- Exploration and Curiosity;
- Empathy and Other-Focus.

Each item presented four response alternatives ordered according to increasing expression of the target behavioural tendency. Rather than identifying correct or incorrect answers, the instrument records individual preferences under standardised hypothetical situations. Consequently, the assessment captures how students tend to approach decision-making within a controlled narrative environment instead of evaluating factual knowledge or academic competence.

Individual item responses were transformed into ordinal scores ranging from one to four. Dimension scores corresponded to the arithmetic mean of the five items composing each behavioural construct, producing continuous indices varying between 1.00 and 4.00. The global behavioural profile was obtained by averaging the eight dimension scores, generating a single summary indicator describing the participant's overall decision pattern within the assessment.

Neither dimensional nor global scores were interpreted as fixed personality traits. Throughout this investigation they are understood as behavioural tendencies expressed under the specific contextual conditions created by the role-playing environment.

## Academic Performance Indicators

To examine the educational relevance of the behavioural profiles, Psychogenesis-RPG scores were analysed together with students' academic records routinely produced by the school.

Academic performance was considered a multidimensional educational outcome rather than a direct measure of cognitive ability. School achievement reflects interactions among prior knowledge, instructional quality, motivation, attendance, study habits, family context, assessment practices, and behavioural self-regulation. Consequently, no assumption was made that behavioural dimensions should individually predict grades.

Instead, academic indicators were used as external educational criteria for investigating whether systematic behavioural differences identified through the RPG corresponded to observable differences in school performance.

This perspective aligns with contemporary educational psychometrics, which emphasises the examination of relationships between behavioural measures and meaningful external educational variables rather than assuming direct causal effects.

## Discussion and Conclusions

From a psychometric perspective, the results provide encouraging evidence that behavioural choices produced in a structured role-playing environment can generate internally coherent dimensions and demonstrate theoretically consistent relationships with external personality measures. The convergent analyses with the age-eligible subsample completing the 16 Personality Factor Questionnaire support the interpretation that selected RPG dimensions capture behavioural tendencies conceptually related to established personality constructs. Importantly, these associations were interpreted as evidence of convergent validity at the level of rank ordering rather than as proof of score equivalence between instruments. Such an approach recognises that the RPG and the 16PF differ fundamentally in measurement philosophy, administration procedures, and intended educational purposes.

This study demonstrates that structured role-playing games may function as psychometrically informed environments for generating behavioural evidence relevant to educational practice. By transforming character-customisation decisions into multidimensional behavioural indicators, the proposed framework extends the educational role of RPGs beyond motivation and engagement toward formative student profiling. The findings suggest that behavioural tendencies expressed through gameplay are associated with meaningful educational differences and may complement conventional academic indicators without replacing established psychometric instruments or professional psychological assessment. As educational systems increasingly seek personalised approaches capable of recognising the complexity of student development, game-based behavioural assessment offers a promising pathway for integrating psychometric rigour, pedagogical relevance, and practical applicability within the everyday reality of Basic Education.

## Acknowledgements

The authors gratefully acknowledge the Coordination for the Improvement of Higher Education Personnel (CAPES), Brazil, for the financial support provided through a doctoral scholarship. This study was financed in part by the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Finance Code 001.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that generative artificial intelligence (ChatGPT, developed by OpenAI) was used solely to assist in translating the original manuscript from Portuguese into English. The use of this technology was strictly limited to language translation and linguistic refinement to ensure clarity, fluency, and conformity with academic English.

The authors confirm that all ideas, study design, methodology, data collection, analyses, interpretations, discussions, and conclusions are entirely their own and were developed through the original conduct of the research. The authors carefully reviewed and edited the translated manuscript and assume full responsibility for the accuracy, integrity, and final content of the submitted work.

## References

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 61–75). Springer. https://doi.org/10.1007/978-1-4614-3305-7_4

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56*(2), 81–105. https://doi.org/10.1037/h0046016

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/BF02310555

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281–302. https://doi.org/10.1037/h0040957

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15). ACM. https://doi.org/10.1145/2181037.2181040

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave Macmillan.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 3025–3034). IEEE. https://doi.org/10.1109/HICSS.2014.377

Kline, P. (1994). *An easy guide to factor analysis*. Routledge.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Mischel, W. (1968). *Personality and assessment*. Wiley.

Moura, W. S., Delbem, E., Dias, A. F. S., & França, J. B. S. (2026). Convergent validation protocol for 16PF–Psychogenesis-RPG based on rank-order convergence and internal consistency. In *Proceedings of the 18th International Conference on Computer Supported Education (CSEDU 2026)* (Vol. 2, pp. 1518–1525). SCITEPRESS. https://doi.org/10.5220/0014803900004021

Moura, W. S., Silva, M. F., Sampaio, J. O., Souza, T. G., Marinho, E. C., & Prado, V. (2021). A social network approach for student's school performance measurement. In *Proceedings of the 13th International Conference on Computer Supported Education (CSEDU 2021)* (Vol. 2, pp. 311–318). SciTePress. https://doi.org/10.5220/0010480003110318

Russell, M. T., & Karol, D. L. (2002). *16PF Fifth Edition administrator's manual*. Institute for Personality and Ability Testing.

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior, 69*, 371–380. https://doi.org/10.1016/j.chb.2016.12.033

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. *Computer Games and Instruction, 55*(2), 503–524.

Shute, V. J., Ventura, M., & Kim, Y. J. (2013). Assessment and learning of qualitative physics in *Newton's Playground*. *The Journal of Educational Research, 106*(6), 423–430. https://doi.org/10.1080/00220671.2013.832970

Siemens, G., & Long, P. (2011). *Penetrating the fog*: Analytics in learning and education. *EDUCAUSE Review, 46*(5), 30–40.

Siqueira Moura, W., Baptista dos Santos França, J., Fonseca da Silva Dias, A., Delbem, E., & Guimarães de Souza, T. (2025). Strategic Alignment of Branches in a Private Educational Network: An Analysis Using Business Process Modeling (BPMN) and Key Performance Indicators (KPIs). *ISYS - Journal of Information Systems, 18*(1), 6:1 – 6:18. https://doi.org/10.5753/isys.2025.5360

Snow, R. E., & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 263–331). American Council on Education/Macmillan.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

# Augmented Reflection in Practicum: AI as a Dialogic Partner in Teacher Education

Anna Jo Perry, Manukau Institute of Technology–Unitec, New Zealand

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Practicum placements are central to field-based qualifications, offering students in person opportunities to develop professional practice. This presentation explores how augmented reflection with AI supported Bachelor of Education (Early Childhood Teaching) students in Aotearoa New Zealand as they navigated practicum expectations across three years of study. Students were assessed against 18 Key Teaching Tasks (KTTs), progressing from beginner to proficient levels. For some, this trajectory has proved challenging. Framed within a post-digital paradigm, the AI functioned as a non-human dialogic partner, prompting, reframing, and scaffolding students' reflective processes. Drawing on survey and focus group data, the study illustrates how the AI surfaced implicit assumptions, deepened pedagogical awareness, and fostered metacognitive insight. Crucially, the AI did not replace human collaboration but complement it, mediating reflection without diminishing learner agency. This hybrid form of reflection created a space in which understandings of practice levels became visible over time. Rather than positioning AI as a sentient collaborator or a mere tool, this work conceptualizes the machine as a post-digital presence, an entangled participant in meaning-making and recursive inquiry. The findings suggest that such configurations can enrich practicum learning by supporting students to articulate professional decisions and navigate complexity with greater confidence.

iafor
The International Academic Forum
www.iafor.org

## Introduction

Reflective practice has become a core part of many disciplines. Schön (1983), drawing on Dewey (1933), described reflection as the way practitioners revisit previous events, recognise their implicit knowledge, and consider how this shapes their actions. In early childhood teacher education, reflective practice is central, appearing in many forms of assessment, including practicum placements.

This paper examines how AI became a support presence in a Bachelor of Education (Early Childhood Teaching) (B.Ed. [ECT]) at an Aotearoa New Zealand institution. A pattern of difficulty passing practicum had emerged, particularly around students' progression through the developmental levels of beginner, intermediate, and proficient. These levels were achieved by the students in their understanding of the programme's eighteen Key Teaching Tasks (KTTs), drawn from the Standards and Code for the Teaching Profession in Aotearoa New Zealand (Teaching Council of Aotearoa New Zealand, 2017). Some students appeared to struggle with understanding and articulating what progression through the levels looked like and how to demonstrate it in practice. Their reflective opportunities were limited to scheduled classroom sessions or centre-based interactions, leaving little space for individualised or on-demand support.

To address this, AI was introduced to provide an additional reflective space. The research aim was to scaffold students' initial interactions with the AI and then observe how they engaged with it independently. The study was guided by the question: "What forms of mediation emerge when AI engages in reflective dialogue with early childhood student teachers during practicum?"

This question focuses on how the AI was seen by the students, what kind of work they believed it would do at the beginning and how it helped to reorganise reflective activity. The core of this engagement was the role it would play rather than being treated as a replacement for human supervision or a simple feedback tool.

The study takes a postdigital perspective, recognising that the practicum learning ecology consists of relationships among multiple presences, both human and nonhuman (Fawns, 2022) Jandrić et al., (2018, p. 895) suggest that the "The postdigital is hard to define; messy; unpredictable; digital and analog; technological and non-technological; biological and informational." The research here shows this through its theoretical groundedness in Vygotsky's (1978) sociocultural theory. In particular, it focuses on the More Knowledgeable Other (MKO) and the Zone of Proximal Development (ZPD), which frame learning as mediated by tools, language, and social interaction.

## Theoretical Framework

This study brings together sociocultural theory and postdigital perspectives to conceptualise AI as a nonhuman mediating presence in reflective practice. This research draws on Vygotsky's (1978) sociocultural theory, which argues that learning is shaped through participation in social and cultural practices and changed through tools and signs. Traditionally, in this theory, Vygotsky distinguished between human mediators (such as an MKO supporting a learner within their ZPD) and using tools in the process, (which were understood as passive mediational artefacts). In this sense, AI functions as a mediational artefact capable of influencing how students organise and extend their thinking.

In this practicum, because of the difficulties some students had with the developmental levels required, the teaching team adopted AI to support understanding for all the students (Jandrić et al., 2018). This nonhuman presence was designed to support knowledge of the KTTs, the different developmental levels and the SMART goals the students had to design for practicum. Importantly, since the inception of the ECE curriculum in 1996 in Aotearoa New Zealand, Early Childhood Education has been built on the tenet that "children learn through responsive and reciprocal relationships with people, places, and things" (Ministry of Education, 2017, p. 21) which also clearly recognises nonhuman presences in this process.

AI is not a human MKO, but still engages interactively, generating prompts, reframing ideas, and responding to learners' language. This positions the machine as a nonhuman mediator of change within the reflective space, capable of shaping how students organise, articulate, and extend their thinking. This research aimed to understand how the students interacted with the AI in terms of reflecting on their practice and supporting them to go deeper in their thinking.

During this research, trouble was taken not to anthropomorphise the AI. Instead it was seen as a nonhuman presence and the students knew and understood this from the beginning.

## Literature Review

There are two interrelated elements to consider within the literature. First, the challenges of engaging in reflection during practicum, and second, the subsequent process of meaning-making that contributes to the development of professional knowledge.

Dewey (1976) described reflection as "an endless mulling over of things… the futile attempt to spin truth out of inner consciousness" (p. 160). While this conceptualisation captures the complex and often uncertain nature of reflection, it also highlights the difficulty of engaging meaningfully in reflective processes. Crandall (2000) suggests that student teachers need to "develop more informed practice, [make] tacit beliefs and practical knowledge explicit … leading to new ways of knowing and articulating" (p. 40). For experienced teachers, this process may be more structured and purposeful but for students undertaking practicum, reflection can remain tentative, descriptive, and unresolved. "From a practical viewpoint, there is no published evidence to show improved teacher or student performance resulting from reflective techniques, the personality of teachers is a missing variable in almost all discussions of reflection" (Akbari, 2007, p. 193).

Building on Dewey's work, Schön (1983) introduced the concepts of reflection-in-action and reflection-on-action, emphasising the temporal and situated nature of reflective thinking within professional practice. These ideas are particularly relevant for practicum settings, where students must navigate immediate decision-making alongside retrospective evaluation of their actions. Similarly, Eraut (2008) argued that reflection occurs both on and off the job and plays a crucial role in recognising and learning from mistakes. Moon (2004) also looked at the importance of being able to apply learning in the messiness of workplaces and the importance and complexity of the learning that came from such situations.

Together, these perspectives highlight key challenges associated with learning through practicum. First, reflection is grounded in personal experience, meaning that the development of professional knowledge is shaped by prior beliefs, values, and experiences (Schön, 1983). Second, despite what writers have posited, these prior experiences may either support or constrain the ability to recognise and learn from mistakes.

In field-based early childhood programmes, practicum itself comprises two key components: regular weekly placements within early childhood centres, and extended block placements, often undertaken in unfamiliar or "away" centres. These experiences are designed to expose students to diverse early childhood age groups, pedagogical approaches, and institutional philosophies. However, the reflective process within these contexts is frequently mediated by mentor teachers, whose own perspectives and experiences also may influence, and at times complicate, students' meaning-making. As a result, students' reflective practice may become shaped not only by their own experiences, but also by the interpretive frames offered by others.

Reflection within practicum settings therefore is not an individual cognitive act, but a distributed cognitive process shaped across people, artefacts, and contexts (Vygotsky, 1978). Reflective knowledge of practice emerges through interactions with mentor teachers, peers, routines, and pedagogical documentation (including assignment requirements), with each contributing distinct expectations and interpretive frames.

Within this distributed landscape, no single actor holds complete authority over meaning-making. Instead, understanding is and has been co-constructed through engagement with multiple sources of knowledge. The introduction of AI into this research further extends this distribution. The AI itself is embedded in the documentation used to guide it and the experiences informing its development, as it becomes another node within the reflective extended network.

This expansion highlights that the "more knowledgeable other" (MKO) is never singular, but in this project is developed through a dynamic grouping of human and nonhuman influences that collectively shape reflective practice.

Although reflective practice has been extensively explored within teacher education, relatively little research has examined how reflective processes are mediated through AI in postdigital practicum contexts. This study contributes to this emerging area by investigating AI as a nonhuman mediating presence within reflective dialogue.

This research is situated within a postdigital paradigm, which recognises that human and digital elements do not operate as separate entities but are already entangled within educational practice. While Vygotsky's original conception of mediation positioned tools as instruments under human control, postdigital perspectives complicate this distinction. The AI in this study does not function as a tool in the traditional sense, nor as a fully autonomous collaborator. Rather, it operates as a nonhuman presence that participates in dialogue and contributes to meaning-making (Jandrić et al., 2018). In this way, the AI can be understood as a dialogic partner that both responds to and reshapes reflective processes, without possessing intentionality or agency in human terms.

While reflection in teacher education has been widely explored, less attention has been given to how reflective processes are mediated within postdigital environments where nonhumans participate in meaning-making. In particular, there remains limited understanding of how AI systems function not simply as tools, but as mediating presences that actively reorganise reflective activity.

This study addresses this gap by examining how AI participates in and reshapes reflective dialogue during practicum, extending current understandings of reflection as a distributed, mediated, and increasingly postdigital process.

## Context and Methodology

The context for this research was a Bachelor of Teaching (Early Childhood Education) degree in Aotearoa New Zealand. Participants were 38 student teachers across two third-year cohorts preparing for practicum placements of between three and five weeks within a 17-week semester.

Practicum assessment was structured around 18 Key Teaching Tasks (KTTs), aligned with the Standards for the Teaching Profession (Teaching Council of Aotearoa New Zealand, 2017). Across the three-year degree, student progression is framed through a developmental hierarchy of beginner, intermediate, and proficient levels. However, internal programme data indicated a small, recurring pattern whereby some students experienced difficulty transitioning between these levels and occasionally resulting in unsuccessful placements.

In response to this identified challenge, the teaching team introduced AI to support students' reflective processes during practicum. The AI was made available to both cohorts via a link embedded within the institutional Learning Management System (LMS).

This study initially adopted a Design-Based Research (DBR) approach, allowing the identification of an educational problem situated in practice and the iterative development of a pedagogical intervention (Wang, & Hannafin, 2005). As the study progressed, the methodology expanded into a mixed methods design. Data were collected through student surveys, and two focus groups of six participants each, representing the two cohorts.

Data were analysed using thematic analysis to identify patterns in how students engaged with reflective processes and the role of the AI in mediating these experiences. The data was read over multiple iterations looking for themes that emerged naturally from the data (Terry & Hayfield, 2021).

## Findings

Earlier, the research question was posed and now is an appropriate time to answer it.

"What forms of mediation emerge when AI engages in reflective dialogue with early childhood student teachers during practicum?"

Analysis of the focus group transcripts and survey suggest three key findings showing mediation in the students' interactions with the AI.

First, students initially positioned the AI as a generic AI tool, but over time described it as a more situated pedagogical presence. At the outset, the AI was framed in familiar, instrumental terms:

"Oh, it's just another AI…." (student response)

However, as students reflected on their ongoing interactions, this positioning shifted through continued mediation by the AI. Rather than as the generic tool from the start of the practicum, the AI was increasingly described in relation to their learning and practice indicating a shift in how they understood the AI's role:

"It gives me examples and theories that I can implement." (Student Response)

"It breaks it down for you." (Student Response)

This suggests that, through use, the AI came to be understood not as abstract technology but as something embedded in the learning context and responsive to the students' needs.

Second, students described the AI as reorganising their reflections through structured re-presentation. Rather than simply as an answer machine adding new content, the students changed their description to how the AI appeared to work with what they provided, extracting, rearranging, and clarifying their ideas. Students emphasised how this made their thinking more accessible and represents a second form of mediation:

"It just takes out the specific information… makes it easier to understand." (Student Response)

"It actually tells us what it means…it's easier now." (Student Response)

"It will give you guidance and kind of steer you." (Student Response)

Finally, interactions with the AI reconfigured students' awareness of their own practice.

Students described the AI as making expectations visible and positioning them in relation to their reflective process. This appeared to support a clearer sense of where they were and what was required, suggesting a reorganisation of their original thinking:

"I wanted help going in depth with my reflection." (Student Response)

"It actually tells you which one I'm going to." (Student Response)

"Without it I was lost… now I feel…." (Student Response)

Clearly in all three themes the AI functions as a mediating presence that reshapes students' explanations, helping them to see their thinking in a more ordered and meaningful way.

In this way, the AI supported students to recognise depth, progression, and purpose within their reflections, enabling them to orient themselves more confidently within their practicum learning.

## Discussion

This study set out to explore how reflection is experienced within a postdigital practicum environment in which AI is present. While practicum can be understood as a distributed learning ecology, the findings suggest that reflective activity is not evenly distributed across all elements. Instead, the AI occupies a distinct mediating role that reshapes how reflection is enacted and understood. This becomes clear in that reflection is not visible as solely an individual cognitive act, but a process that becomes newly visible and reorganised through interaction with nonhuman presence.

Foundational conceptions of reflective practice, such as those proposed by Dewey (1933) and Schön (1983), position reflection primarily as an individual cognitive process that occurs after

or during action. However, in practice, reflection in practicum has always been relationally and materially distributed across mentors, institutional expectations, artefacts, and prior experience. Rather than suggesting these theorists were incorrect, it is more accurate to view their work as foregrounding the individual while leaving the broader relational conditions of reflection less explicit.

From this perspective, sociocultural theory (Vygotsky, 1978) provides a more distributed account of mediation, where learning is shaped through interaction with tools and more knowledgeable others (MKOs). However, the findings of this study extend this framing further. The AI does not function solely as a passive tool nor as a human-like MKO. Instead, it introduces a form of responsive mediation that participates in the organisation of reflective thinking itself. This could be seen as a shift in mediation from supporting reflection to actively shaping its form.

More-than-human and postdigital perspectives (Fawns, 2022) help clarify this shift. In a postdigital paradigm, human and technological elements are not separate but already entangled within educational practice. Within this framing, the AI is best understood not as an external addition to reflection but as part of the "reflective process assembly." However, unlike other mediating artefacts, the AI introduces a form of dialogic responsiveness that helps to reorganise student thinking in real time. This means reflection is not only distributed across actors but also dynamically reconfigured through interaction.

The findings of this research suggest that reflection in this assembly is not simply retrospective or internal, but continuously reconstituted through interaction with the AI. This compresses the distinction between reflection-in-action and reflection-on-action (Schön), as meaning is shaped in close temporal proximity to the experience rather than after it.

From a sociocultural perspective, this extends Vygotsky's (1978) notion of mediation. While tools and MKOs traditionally support learning processes, the AI observed here plays a more active mediating role. It does not simply assist reflection, it reorganises the conditions under which reflection becomes possible and intelligible.

This is where postdigital entanglement becomes important. The AI's role cannot be reduced to either design features or human interpretation alone. Instead, reflective practice emerges through the entanglement of students, institutional frameworks, practicum expectations, and AI responsiveness. Within this assembly, different actors contribute in qualitatively different ways, but none, importantly, operate independently.

However, this does not position AI as equivalent to human educators. Rather, it highlights that AI introduces a distinct form of mediation that is structured, responsive, and non-evaluative, yet capable of shaping attention and interpretation. This combination differentiates it from both human dialogue and traditional tools.

Overall, here, reflection is seen as a postdigital, entangled practice in which meaning is co-constructed through interaction between human and nonhuman actors. The AI does not simply support reflection, it participates in shaping and deepening what reflection becomes. Recognising this, shifts reflection from an individual cognitive activity to an emergent process of reconfiguration within a distributed learning ecology.

## Limitations of the Research

This study has several limitations that should be acknowledged. First, the findings are based on a small-scale qualitative study and are therefore not intended to be generalisable. The value of the work lies in its exploratory and conceptual insights rather than in claims of representativeness. Second, the study draws solely on students' retrospective accounts of their experiences, rather than on the interactional data itself. While this aligns with the study's focus on how reflection was experienced by students, it limits the ability to examine the micro-processes through which mediation occurred.

## Conclusion

This study has argued that reflective practice within practicum can no longer be understood solely as an individual or even human-centred activity. Instead, it emerges through a postdigital assemblage in which human and nonhuman presences are entangled in the ongoing production of meaning. The AI examined here did not act as a replacement for existing pedagogical relationships, but as a distinct mediating presence that reorganised how reflection was experienced, articulated, and understood. In doing so, it made visible the conditions under which reflective practice becomes possible, particularly for students navigating complex expectations of progression. While modest in scale, this work contributes to a growing recognition that reflection in teacher education is being reshaped within postdigital contexts, calling for a reconsideration of how mediation, agency, and learning are conceptualised in practicum settings.

## References

Akbari, R. (2007). Reflections on reflection: A critical appraisal of reflective practices in L2 teacher education. *System, 35*(2), 192–207.

Crandall, J. (2000). Language teacher education. *Annual Review of Applied Linguistics, 20*, 34–55.

Dewey, J. (1933). *How we think*. Houghton Mifflin.

Dewey, J., & Boyston, J. (1976). *The collected works of J. Dewey: The middle works 1899-1924, vol 1, 1899-1901*. Southern Illinois University Press.

Eraut, M. (2008). *Using research into how professionals learn at work for enhancing placement learning*. WACE Asia Pacific Conference.

Fawns, T. (2022). An entangled pedagogy: looking beyond the pedagogy-technology dichotomy. *Postdigital Science and Education, 4*(3), 711–728.

Jandrić, Knox, Besley, Ryberg, Suoranta & Hayes. (2018). Postdigital science and education. *Educational Philosophy and Theory, 50*(10), 893–899, https://doi.org/10.1080/00131857.2018.1454000

Ministry of Education. (2017). *Te Whāriki: Early Childhood Curriculum*. He whāriki mātauranga mō ngā mokopuna o Aotearoa.

Moon, J. (2004). *A Handbook of Reflective and Experiential Learning: Theory and Practice*. Routledge.

Schön, D. A. (1983). *The reflective practitioner: how professionals think in action*. Basic Books.

Teaching Council of Aotearoa New Zealand. (2017). *Our Code, Our Standards. Code of Professional Responsibility and Standards for the Teaching Profession*. Nga Tikanga Matatika mo te Haepapa Ngaiotanga me nga Paerewa mo te Umanga Whakaakoranga.

Terry, G., & Hayfield, N. (2021). *Essentials of Thematic Analysis*. American Psychological Association.

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research & Development, 53*, 5–23.

# Maker Culture and Sustainability in Basic Education

Lígia Maura Delbem, VespeR, Brazil
Waldir Siqueira Moura, Federal University of Rio de Janeiro, Brazil
Edgar Delbem, Federal University of Rio de Janeiro, Brazil
Célia Regina Tomachuk dos Santos Catuogno, University of São Paulo, Brazil

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This paper examines the integration of the Maker Culture into the pedagogical practices of teachers working in the early years of primary education as a strategy for developing socio-environmental competencies aligned with the Brazilian National Common Core Curriculum (BNCC) and the Sustainable Development Goals (SDGs). Adopting a qualitative approach, the study is designed as a case study with action research characteristics and was conducted in a philanthropic school located in the municipality of Tremembé, São Paulo, Brazil, involving five teachers and 94 students. The theoretical framework is grounded in Critical Environmental Education, active learning methodologies, the Maker Culture, and the Educational Psychogenetics proposed by Fr. Joaquim Ferreira Xavier Jr, sjr. The investigation combined a literature review, questionnaires and interviews, teacher training workshops, pedagogical monitoring, and the implementation of sustainable maker practices. The findings indicate that, although teachers recognised the importance of sustainability, their initial practices were isolated and insufficiently integrated into the curriculum. Following the training process, greater interdisciplinarity, increased student agency, and significant changes in socio-environmental attitudes were observed. The study concludes that the Maker Culture, when integrated with Environmental Education and Educational Psychogenetics, represents a promising pathway for pedagogical innovation and the development of critical citizenship in Basic Education.

*Keywords:* Maker Culture, sustainability, environmental education, teacher education, educational psychogenetics

iafor
The International Academic Forum
www.iafor.org

## Introduction

The contemporary environmental crisis poses complex challenges to school education, requiring pedagogical approaches capable of integrating scientific knowledge, ethical development, and social responsibility. In Brazil, the National Environmental Education Policy (Law No. 9,795/1999) establishes Environmental Education as a continuous, integrated, and cross-curricular practice across all levels of education (Brasil, 1999). In this context, the Brazilian National Common Core Curriculum (BNCC) reaffirms the role of schools in fostering critical, responsible, and sustainability-oriented citizens (Brasil, 2018).

Maker Culture has emerged as a pedagogical approach capable of addressing these challenges by emphasising learning through making, authentic problem-solving, collaboration, and creativity (Bacich & Moran, 2017; Valente, 2018). By bridging theory and practice, Maker Education promotes meaningful learning experiences, particularly in the early years of primary education, a critical stage for the development of values, attitudes, and lifelong habits.

This study is also grounded in the Educational Psychogenetics developed by Fr. Joaquim Ferreira Xavier JR, whose work highlights the role of proprioceptive experiences, concrete action, and subject–object interaction in the construction of knowledge (Xavier Jr., 2006). According to the author, cognitive and ethical development emerges through lived experience, experimentation, and the learner's active engagement with the surrounding world—principles that closely align with the Maker approach.

The aim of this paper is to analyse how the integration of Maker Culture into teachers' pedagogical practices contributes to the development of socio-environmental competencies and to the consolidation of sustainable practices in Basic Education.

## Theoretical Framework

Maker Culture originated in the *Do It Yourself* (DIY) movement, which gained momentum through the widespread availability of accessible technologies, the growth of the sharing culture, and the appreciation of collaborative creation (Anderson, 2012; Zylbersztajn, 2015). Historically, DIY extends beyond a technical practice and represents a cultural attitude centred on authorship: repairing, adapting, building, and recreating become actions that shift individuals from the role of passive consumers to active producers of meaning and solutions. As access to digital resources and open-source culture expanded, this logic of authorship intensified and became institutionalised within networks, communities, and collaborative environments such as makerspaces and hackerspaces, where knowledge is generated through experimentation, trial and error, the sharing of prototypes, and collective refinement (Zylbersztajn, 2015). Anderson (2012) interprets this movement as part of a broader transformation in contemporary production, whereby tools once restricted to industrial settings have become locally accessible, fostering distributed innovation and solutions driven by real-world needs.

Within education, this approach challenges the traditional transmission-based model by promoting practices centred on the learner, inquiry, and experimentation (Papert, 1994). Learning is understood less as the passive reception of information and more as an active process of creation: producing tangible artefacts, testing hypotheses, observing outcomes, and refining solutions transform the classroom into a space of investigation and authorship. From this perspective, Papert's constructionism argues that thinking develops most effectively when

learners create meaningful artefacts that externalise their reasoning processes, thereby fostering metacognition and conceptual refinement (Papert, 1994). From a curricular perspective, this implies reorganising educational objectives, learning time, and teaching methods so that "making" becomes the central organising principle of learning. Projects become the primary unit of educational experience, integrating knowledge, procedures, and values around authentic, context-based problems.

At this point, the contribution of Fr. Joaquim Ferreira Xavier Jr, sjr. provides an important conceptual foundation for interpreting Maker Culture in the early years of primary education. Through the development of Educational Psychogenetics, Xavier argues that children's cognitive activity is primarily logical-concrete, relying on manipulation, interaction with physical objects, and sensorimotor experiences for the consolidation of coherent cognitive structures (Xavier Jr., 2006). This perspective shifts learning from the abstract domain to that of guided experience, where learning consists of performing actions, recognising patterns, constructing relationships, and reorganising psychological structures through processes of reversal and reversibility—that is, operations progressively refined through practice, comparison, and reflection upon one's own actions (Xavier Jr., 2006). Consequently, Maker Culture does not merely "fit" childhood because it is engaging or enjoyable; rather, it responds to an epistemological requirement of child development. Children between six and ten years of age learn more effectively when they are able to experiment, manipulate materials, make mistakes, reconstruct ideas, and integrate their experiences into increasingly stable cognitive schemas (Xavier Jr., 2006). Within this framework, curiosity—commonly regarded as the driving force of Maker Education—can be understood as both an attentional and affective condition for learning: novelty evokes emotion, and emotion enhances attention, facilitating knowledge retention and cognitive reorganisation (Xavier Jr., 2006).

According to Bacich and Moran (2017), active learning methodologies promote greater engagement and meaningful learning by involving students in contextualised problem situations that require participation, collaboration, argumentation, and decision-making. Maker Culture operates through practical challenges and design-based problems that encourage learners to conceive, test, and validate solutions while integrating critical thinking and creativity. Valente (2018) further argues that these practices are aligned with the competencies required in the twenty-first century, as they simultaneously mobilise cognitive, intrapersonal, and interpersonal dimensions, fostering collaboration, problem-solving, and autonomy. From Xavier's perspective, however, autonomy should not be regarded merely as an abstract educational objective but rather as a progressive achievement grounded in experience. Through active engagement with the surrounding world and with material objects, learners expand their capacity for action, develop awareness of their own learning processes, and acquire the confidence necessary to undertake intellectual and practical challenges (Xavier Jr., 2006). Consequently, Maker Culture may be understood as a pedagogical environment that supports the maturation of psychological mobilisation and cognitive elaboration because it provides concrete situations of purposeful action guided by explicit goals and evaluative criteria (Xavier Jr., 2006).

When integrated with Environmental Education, Maker Culture enhances critical reflection on the use of natural resources, responsible consumption, and socio-environmental responsibility (Carvalho, 2012; Jacobi, 2003). This occurs because maker activities, through their emphasis on materials, waste, reuse, and prototyping, make visible aspects of consumption that often remain unnoticed, including resource origins, material life cycles, and environmental impacts. Environmental Education, conceived as a political and pedagogical process, seeks to develop

individuals capable of understanding the complexity of socio-environmental issues and participating collectively in the search for sustainable solutions (Jacobi, 2003). Maker experiences provide an appropriate context for integrating reflection and action, enabling students to experience practices such as reduction, reuse, and recycling within authentic school projects rather than treating these concepts merely as theoretical principles. Carvalho (2012) argues that schools can foster environmental ethics through dialogue among different forms of knowledge and collaborative action. By organising learning around teamwork and co-creation, Maker Culture strengthens this collaborative and community-oriented dimension.

The United Nations 2030 Agenda further reinforces the role of education as a fundamental pillar of sustainable development, particularly through Sustainable Development Goal 4 (Quality Education) (UN, 2015; UNESCO, 2017). UNESCO (2017) emphasises that education for the Sustainable Development Goals should foster competencies such as critical thinking, collaboration, empathy, and the capacity to address complex problems. Sustainable maker projects connect global challenges with local realities, promoting situated citizenship by enabling students to recognise that everyday decisions—such as water consumption, waste management, and material reuse—have tangible consequences and can be transformed through conscious and collective action. Furthermore, Maker Culture encourages interdisciplinarity by integrating science, mathematics, language, and the arts in the creation of objects, campaigns, prototypes, and narratives, thereby narrowing the gap between curriculum and everyday life.

Within this framework, Xavier's research broadens the educational scope of these practices by explaining why learning for sustainability must be experiential rather than merely informational. If, as proposed by Xavier, the human capacity for innovation is supported by the learner's plasticity and progressive reorganisation when confronted with authentic challenges, then education for sustainability becomes coherent when structured around experiences that engage the body, perception, action, and reflection (Xavier Jr., 2006). Sustainability thus becomes simultaneously content, method, and ethos: it is not simply a topic to be studied but a way of acting and making decisions, developed through practice, intentional pedagogical mediation, and continuous opportunities for revision and improvement. Consequently, the articulation between Maker Culture, Environmental Education, and the 2030 Agenda may be understood as a comprehensive pedagogical strategy for holistic education. By learning through making, students develop cognitive, social, and ethical competencies while internalising habits and values grounded in responsibility and care for the shared world (UN, 2015; UNESCO, 2017; Xavier Jr., 2006).

## Methodology

This study adopts a qualitative applied research approach and is designed as a case study with characteristics of action research (Marconi & Lakatos, 2010; Thiollent, 2011). The research was conducted in a philanthropic school located in Tremembé, São Paulo, Brazil, serving students in the early years of primary education. The school is situated in a predominantly rural area characterised by social vulnerability. It was established through the *Breaking the Cycle of Poverty* project, conceived by Fr. Joaquim Ferreira Xavier Jr, sjr., whose educational philosophy emphasises holistic development and education as a means of social emancipation (Xavier Jr., 2006). This institutional context is considered not merely as the setting of the investigation but as an interpretative component of the study, since the school's culture, community values, and educational identity directly influence the adoption of sustainable practices and active learning methodologies, particularly Maker Culture.

The participants comprised five classroom teachers (one representing each grade from Year 1 to Year 5) and 94 enrolled students. However, only 62 students participated effectively in the empirical phase after formal parental consent had been obtained, a prerequisite for research involving minors. This difference between the total student population and the effective research sample is methodologically significant because it defines the ethical boundaries of the study and delimits the corpus of analysed data. Teacher participation was voluntary and formalised through informed consent forms, ensuring confidentiality, anonymisation of data, and the right to withdraw from the study at any stage, in accordance with established ethical principles for educational research.

The first stage consisted of presenting the project to the school leadership, pedagogical coordinators, and teachers during a collective pedagogical planning session. This introductory phase adopted a dialogical and participatory format and served both diagnostic and relational purposes. It sought to clarify the objectives, theoretical foundations, and implementation schedule while identifying participants' expectations, potential resistance, and practical conditions for implementation. Within the action research tradition, this stage is essential because it establishes a collaborative partnership and promotes shared responsibility between the researcher and participants (Thiollent, 2011).

Subsequently, a diagnostic assessment of teachers' conceptions regarding sustainability, Environmental Education, and Maker Culture was conducted through digital questionnaires containing both closed-ended and open-ended questions. The instrument was designed to assess participants' familiarity with the principles of the 3Rs (Reduce, Reuse, and Recycle) and to examine how sustainability was understood and integrated into classroom practice. The use of an online questionnaire provided flexibility, minimised disruption to school routines, and facilitated subsequent organisation and analysis of the collected data. Analytically, this diagnostic phase fulfilled two complementary functions: (a) generating evidence regarding teachers' prior beliefs, knowledge, and pedagogical practices; and (b) informing the design of the subsequent training programme by identifying observed gaps and emerging professional development needs.

To complement these data, individual semi-structured interviews were conducted with each participating teacher. These interviews sought to deepen the understanding of pedagogical practices, perceived barriers—including time constraints, limited resources, lesson planning demands, and curriculum alignment—and the opportunities for integrating Maker Culture with sustainability education.

Following the initial data collection, questionnaire responses and interview transcripts were organised and analysed through content analysis, following the procedures proposed by Bardin (2011), including pre-analysis, material exploration, and interpretation of findings. This analytical approach enabled the identification of recurring patterns within teachers' narratives and the construction of thematic categories that guided the intervention. The principal themes identified included limited conceptual understanding of Maker Education, a tendency to reduce sustainability to recycling practices, structural constraints affecting implementation, and the demand for continuing professional development.

The intervention phase consisted of teacher training workshops and pedagogical meetings designed to strengthen theoretical foundations and expand teachers' instructional repertoire. The training programme focused on integrating sustainability, the principles of the 3Rs, and interdisciplinary teaching with active learning methodologies and project-based learning.

Within applied research, professional development is not regarded as a peripheral activity but rather as the central mechanism for promoting pedagogical transformation, which is fully consistent with the action research paradigm, where investigation and intervention occur simultaneously (Thiollent, 2011).

Following the training programme, Maker-based activities were implemented in the participating classrooms under continuous pedagogical supervision. Classroom observations, implementation records, and reflective field notes documented the development of the activities throughout the intervention. Lesson planning adopted an interdisciplinary approach, integrating Science, Mathematics, Portuguese Language, and Arts around the themes of the 3Rs and the Sustainable Development Goals (SDGs). In practice, activities included measuring and recording water and electricity consumption (Mathematics); investigating decomposition processes, waste management, and material reuse (Science); producing argumentative texts and environmental awareness campaigns (Portuguese Language); and designing artistic and functional artefacts using recyclable materials (Arts). Within this framework, interdisciplinarity extended beyond the simple juxtaposition of curricular content, instead organising different disciplinary perspectives around a shared problem—sustainability—thereby fostering meaningful curricular integration.

Monitoring and evaluation were conducted systematically and collaboratively through regular meetings, feedback sessions, and progressive adjustments throughout the intervention. This iterative process is characteristic of action research, as it enables continuous feedback between planning, implementation, and reflection, allowing methodological refinements to be introduced throughout the study (Thiollent, 2011). The extensive use of recyclable and reusable materials—including cardboard, plastic bottles, and paper—combined with digital technologies such as online research tools, collaborative digital platforms, and audiovisual resources, reinforced the maker-oriented character of the intervention while expanding opportunities for documentation, communication, and dissemination of students' projects (Valente, 2018).

**Table 1**
*Research Design and Stages of the Intervention*

| Stage | Purpose | Participants | Instruments and Procedures | Data/Outputs Generated | Expected Outcomes |
|---|---|---|---|---|---|
| **1. Institutional awareness and alignment** | To present the project, establish collaborative agreements, and assess implementation conditions | School leadership, pedagogical coordination, and five teachers | Dialogic presentation meeting; definition of the implementation schedule and operational agreements | Records of institutional alignment, participant commitment, and contextual conditions | Formal agreement for participation and a realistic implementation plan |
| **2. Initial diagnosis of teachers' conceptions** | To identify teachers' conceptions of sustainability, Environmental Education, and Maker Culture, as well as their previous pedagogical practices | Five teachers | Digital questionnaire comprising closed-ended and open-ended items | Baseline data on teachers' perceptions, knowledge, and existing practices | Identification of professional development needs and priority themes |
| **3. Qualitative exploration** | To understand pedagogical rationales, perceived barriers, and opportunities for curriculum integration | Five teachers | Individual semi-structured interviews; recorded responses | Interview transcripts and thematic meaning units | Expanded interpretation of the diagnostic findings and contextualisation of responses |

| | | | | | |
|---|---|---|---|---|---|
| **4. Data analysis and categorisation** | To organise and interpret the collected data in order to guide the intervention and subsequent monitoring | Researcher (with feedback provided to participants) | Content analysis following Bardin (2011): pre-analysis, coding, categorisation, and inference | Thematic categories (e.g., sustainability limited to recycling; time and infrastructure constraints; professional development needs) | Analytical foundation for designing the training programme and monitoring criteria |
| **5. Professional development intervention** | To broaden teachers' knowledge of Maker Culture, active learning methodologies, the 3Rs, and the SDGs, while supporting interdisciplinary planning | Five teachers | Teacher training workshops, guided study sessions, project-based planning, and formative feedback | Lesson plans and sustainable Maker project designs | Increased teachers' conceptual understanding and methodological confidence for implementation |
| **6. Classroom implementation** | To implement interdisciplinary Maker activities related to the 3Rs and the Sustainable Development Goals | Five teachers and 62 participating students (from a total enrolment of 94) | Maker projects, inquiry-based activities, artefact construction, and the use of recyclable and reusable materials | Artefacts and prototypes, implementation records, and evidence of student learning | Greater student agency, active learning, and effective interdisciplinary curriculum integration |
| **7. Ongoing pedagogical monitoring** | To monitor implementation, refine strategies, and consolidate sustainable educational practices | Teachers, students, and the pedagogical coordination team | Classroom observations, follow-up meetings, formative feedback, and iterative adjustments characteristic of action research | Monitoring records and documentation of implemented changes | Continuous improvement of pedagogical practices through reflective feedback cycles |
| **8. Dissemination and project culmination** | To validate and communicate project outcomes while strengthening a culture of sustainability within the school | School community | School exhibition or fair, student presentations, and display of projects and reflective activities | Public presentations and records of community participation | Consolidation of a sustainable Maker culture and dissemination of good practices within the institution |
| **9. Educational product** | To systematise guidelines and teaching materials to support replication and scalability | Teachers (primary audience) and the wider educational community | Development of an educational e-book containing teaching sequences, implementation guidelines, and project examples | Pedagogical e-book | Practical resource supporting the implementation of sustainable Maker methodologies in other educational contexts |

## Results and Discussion

The initial findings revealed that the participating teachers possessed only a partial understanding of sustainability, frequently associating it exclusively with recycling practices. This tendency reflects a limited and functional conception of sustainability, centred primarily on waste management and isolated environmental actions, without encompassing the broader ethical, social, economic, and political dimensions of sustainable development. Such findings are common in educational contexts where Environmental Education is treated as a cross-curricular theme but is insufficiently integrated into pedagogical planning, resulting in fragmented activities disconnected from a systemic understanding of sustainability.

By conceptualising sustainability as a multidimensional construct, Sachs (2008) argues that sustainable development requires the integration of environmental, social, economic, cultural, and political dimensions, thereby overcoming reductionist approaches that simplify its complexity into isolated practices. Similarly, Morin (2015) emphasises that contemporary problems cannot be adequately understood through disciplinary fragmentation but instead require complex thinking, capable of recognising relationships, interdependencies, and both short- and long-term consequences. The initial diagnosis, which demonstrated that sustainability was predominantly equated with recycling, therefore reveals both a conceptual gap and a pedagogical limitation. When teachers approach sustainability from such a restricted perspective, classroom activities tend to remain disconnected from both the curriculum and students' lived experiences, reducing their educational and transformative potential.

The professional development workshops contributed significantly to broadening teachers' understanding by incorporating the social, economic, and ethical dimensions of sustainability. However, this transformation should not be interpreted merely as the acquisition of additional theoretical knowledge. Rather, it represented a reconfiguration of teachers' pedagogical perspectives, enabling sustainability to be understood as an organising principle for curriculum planning rather than as an isolated supplementary topic. Following the training programme, teachers demonstrated a greater tendency to redesign their instructional planning with increased intentionality, selecting authentic problems, valuing the local context, and integrating competencies established by the Brazilian National Common Core Curriculum (BNCC), particularly those related to citizenship, values, and responsible action. The training programme also functioned as a legitimising mechanism by providing participants with a shared conceptual language, practical examples, and pedagogical models, thereby reducing uncertainty and increasing teachers' confidence in adopting active learning methodologies.

Maker-based activities—including the construction of reusable artefacts, environmental inquiry projects, and initiatives centred on the reuse of recyclable materials—generated high levels of student engagement, strengthened interdisciplinary learning, and promoted observable behavioural changes both within the school environment and in students' family contexts. These findings may be interpreted from three complementary perspectives.

The first concerns student engagement. By organising learning around authentic challenges, Maker Culture shifts education from a model centred on knowledge reproduction towards one focused on knowledge production. This transition tends to enhance students' motivation, perseverance, and sense of purpose by positioning them as active participants in the learning process rather than passive recipients of information.

The second perspective relates to interdisciplinarity. Because Maker projects are structured around authentic problems, they naturally require the integration of multiple disciplinary perspectives. The development of a single artefact or the investigation of a real-world problem simultaneously involves mathematical reasoning, scientific inquiry, written communication, argumentation, artistic expression, and technological skills, thereby fostering meaningful curricular integration rather than the simple juxtaposition of subject content.

The third perspective concerns behavioural change. When sustainability is experienced through concrete practices—such as reducing waste, reusing materials, redesigning discarded objects, and critically reflecting on consumption—students are more likely to internalise sustainable habits through repeated meaningful experiences, emotional engagement, and social validation

within their peer groups. Consequently, these behaviours tend to extend beyond the school environment and become incorporated into everyday family practices.

These findings are consistent with previous research demonstrating the effectiveness of active learning methodologies in promoting meaningful learning and strengthening student agency (Moran, 2018; Vickery, 2016).

From the perspective of Educational Psychogenetics, these outcomes acquire additional explanatory significance. According to Xavier (2006), cognitive development is consolidated through concrete action, sensory experience, and reflective interaction with the environment. The Maker activities implemented in this study created conditions in which students were able to manipulate materials, formulate hypotheses, test solutions, evaluate outcomes, and reconstruct their ideas through iterative processes. Consequently, sustainability was not approached merely as conceptual knowledge but as lived experience, allowing environmental values to become progressively integrated into students' cognitive and behavioural structures. This interpretation suggests that the observed behavioural changes were not simply the result of increased environmental awareness but reflected deeper processes of cognitive and ethical development supported by meaningful action.

Taken together, the results indicate that integrating Maker Culture with Environmental Education and Educational Psychogenetics creates a pedagogical environment capable of promoting not only conceptual learning but also the development of socio-environmental competencies, interdisciplinary thinking, and responsible citizenship. In this context, sustainability emerges simultaneously as educational content, pedagogical methodology, and ethical orientation, reinforcing the potential of Maker Education to contribute to transformative educational practices in Basic Education.

**Table 2**

*Comparison of Educational Practices Before and After the Intervention*

| Analysed Dimension | Before the Intervention | After the Intervention |
|---|---|---|
| **Conception of Sustainability** | Predominantly limited understanding, mainly associated with recycling and isolated environmental actions. | Broader understanding incorporating social, economic, and ethical dimensions; sustainability recognised as a curricular and educational organising principle. |
| **Curriculum Integration** | Sporadic and poorly systematised activities, with limited integration across curriculum subjects. | More intentional and interdisciplinary planning, integrating sustainability into Science, Mathematics, Portuguese Language, and Arts. |
| **Pedagogical Planning** | Emphasis on isolated activities focused on immediate products and rapid execution. | Emphasis on project-based learning, sequential processes, documentation, inquiry, reflection, and continuous assessment. |
| **Teacher's Role** | Predominantly transmissive and instructional, with limited opportunities for experimentation. | Facilitator and inquiry-oriented role, supporting decision-making, collaboration, and problem-solving throughout the projects. |
| **Student Engagement** | Variable engagement, often dependent on occasional “different” activities lacking continuity. | High and sustained engagement driven by authentic challenges, resulting in greater autonomy, participation, and sense of purpose. |
| **Interdisciplinarity** | Fragmented organisation by subject area, with limited connections among curricular contents. | Integration through problem- and project-based learning, mobilising multiple disciplines, languages, and procedures around sustainability challenges. |

| | | |
|---|---|---|
| **Type of Learning** | Predominantly reproductive learning and activities involving limited investigative complexity. | Active and meaningful learning characterised by experimentation, prototyping, inquiry, and argumentation. |
| **Materials and Resources** | Limited use of educational materials, with little emphasis on reuse or prototyping. | Extensive use of recyclable and reusable materials, with reuse practices functioning simultaneously as content and pedagogical methodology based on the **3Rs**. |
| **Socio-environmental Values and Attitudes** | Sustainability approached primarily at the discursive level, with limited evidence of behavioural change. | Internalisation of environmental values through lived experiences, reflected in observable behavioural changes at school and reported transfer of sustainable practices to students' homes. |
| **School Culture** | Sustainability treated as a peripheral cross-curricular theme with limited presence in everyday school life. | Sustainability incorporated into school projects and events, demonstrating potential for long-term cultural consolidation and replication. |

At the same time, from the perspective of Educational Psychogenetics, the findings indicate that concrete action and experimentation facilitated the internalisation of socio-environmental values, reinforcing the central role of lived experience in the educational process (Xavier JR., 2006). Xavier's theoretical contribution helps explain why the intervention extended beyond simply increasing students' interest, promoting more profound cognitive and ethical transformations. If cognitive and moral development is grounded in experiential engagement and the progressive reorganisation of psychological structures through action in the world, then sustainable Maker activities constitute a particularly favourable educational environment. Students learn by manipulating materials, recognising relationships, testing hypotheses, and evaluating the consequences of their decisions.

Consequently, socio-environmental values are not acquired merely as abstract concepts but are internalised through lived experience. Caring for materials, recognising waste, negotiating solutions collaboratively, and understanding the impact of individual and collective choices become processes through which practical awareness is constructed, fostering an ethics embedded in everyday action. Learning, therefore, extends beyond the acquisition of knowledge to encompass the reorganisation of habits, perceptions, and attitudes towards the surrounding world.

The intervention also generated important effects on teachers' professional practice. Through planning, implementing, and monitoring Maker projects, teachers progressively adopted a more inquiry-oriented pedagogical role characterised by observation, documentation, questioning, reflective dialogue, support for student decision-making, and process-oriented assessment. This transformation is consistent with the pedagogical shift promoted by active learning methodologies, in which teachers move beyond the traditional role of knowledge transmitters to become facilitators of learning experiences and curators of students' learning pathways (Moran, 2018). In practice, this change contributed to the development of a learning culture in which mistakes were recognised as an essential part of the learning process, collaboration became a central pedagogical principle, and educational outcomes were evaluated not only through the final products but also through students' reasoning, documentation, reflective processes, and collaborative interactions.

## Conclusions

The findings of this study demonstrate that the integration of Maker Culture, Environmental Education, and Educational Psychogenetics constitutes a coherent and promising approach to

pedagogical innovation in Basic Education. The articulation of these perspectives enabled the study to move beyond fragmented approaches to sustainability, traditionally limited to isolated environmental activities, by promoting pedagogical practices grounded in experience, inquiry, and student authorship. The results confirm that learning is strengthened when students are actively engaged in solving authentic problems, producing artefacts, and reflecting critically upon their own actions, in accordance with the principles established by the Brazilian National Common Core Curriculum (BNCC) and the United Nations 2030 Agenda for Sustainable Development.

The educational legacy of Fr. Joaquim Ferreira Xavier, sjr. proved particularly valuable in interpreting the outcomes of the intervention. Educational Psychogenetics provided a theoretical framework that emphasises the centrality of concrete experience, embodiment, and affective-cognitive mobilisation within the learning process. By demonstrating that socio-environmental values are constructed through lived experience rather than through the simple transmission of information, this research highlights both the contemporary relevance and the enduring significance of Xavier's work for addressing current educational challenges, particularly those related to ethical development and citizenship education during the early years of primary schooling.

## Acknowledgements

The authors gratefully acknowledge the Coordination for the Improvement of Higher Education Personnel (CAPES), Brazil, for the financial support provided through a doctoral scholarship. This study was financed in part by the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Finance Code 001.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that generative artificial intelligence (ChatGPT, developed by OpenAI) was used solely to assist in translating the original manuscript from Portuguese into English. The use of this technology was strictly limited to language translation and linguistic refinement to ensure clarity, fluency, and conformity with academic English. The authors confirm that all ideas, study design, methodology, data collection, analyses, interpretations, discussions, and conclusions are entirely their own and were developed through the original conduct of the research. The authors carefully reviewed and edited the translated manuscript and assume full responsibility for the accuracy, integrity, and final content of the submitted work.

## References

Anderson, C. (2012). *Makers*: *The new industrial revolution*. Crown Business.

Bacich, L., & Moran, J. (2017). *Metodologias ativas para uma educação inovadora*: Uma abordagem teórico-prática [Active methodologies for innovative education: A theoretical and practical approach]. Penso.

Bardin, L. (2011). *Análise de conteúdo* [Content analysis] (Revised and expanded ed.). Edições 70.

Brasil. (1999). *Lei nº 9.795, de 27 de abril de 1999*: Dispõe sobre a educação ambiental e institui a Política Nacional de Educação Ambiental [Law No. 9,795 of April 27, 1999: Establishes Environmental Education and institutes the National Environmental Education Policy]. Diário Oficial da União.

Brasil. (2018). *Base nacional comum curricular* [Brazilian national common core curriculum]. Ministério da Educação.

Carvalho, I. C. M. (2012). *Educação ambiental*: *A formação do sujeito ecológico* [Environmental education: The formation of the ecological subject] (6th ed.). Cortez.

Jacobi, P. R. (2003). Educação ambiental, cidadania e sustentabilidade [Environmental education, citizenship and sustainability]. *Cadernos de Pesquisa, 118*, 189–205. https://doi.org/10.1590/S0100-15742003000100009

Marconi, M. A., & Lakatos, E. M. (2010). *Fundamentos de metodologia científica* [Fundamentals of scientific methodology] (7th ed.). Atlas.

Moran, J. (2018). *A educação que desejamos*: *Novos desafios e como chegar lá* [The education we want: New challenges and how to achieve them]. Papirus.

Morin, E. (2015). *Ensinar a viver*: *Manifesto para mudar a educação* [Teaching to live: Manifesto for changing education]. Sulina.

Organização das Nações Unidas. (2015). *Transformando nosso mundo*: *A Agenda 2030 para o desenvolvimento sustentável* [Transforming our world: The 2030 Agenda for Sustainable Development]. United Nations.

Papert, S. (1994). *A máquina das crianças*: *Repensando a escola na era da informática* [The children's machine: Rethinking school in the age of the computer]. Artmed.

Sachs, I. (2008). *Caminhos para o desenvolvimento sustentável* [Paths to sustainable development]. Garamond.

Thiollent, M. (2011). *Metodologia da pesquisa-ação* [Action research methodology] (18th ed.). Cortez.

UNESCO. (2017). *Education for sustainable development goals*: *Learning objectives*. UNESCO Publishing.

Valente, J. A. (2018). *Inovação nos processos de ensino e aprendizagem*: *O papel das tecnologias digitais* [Innovation in teaching and learning processes: The role of digital technologies]. NIED/UNICAMP.

Vickery, A. (2016). *Creative learning*: *Supporting student creativity through assessment for learning*. Routledge.

Xavier, J. F. (2006). *Psicogenética educacional* [Educational psychogenetics]. VespeR Editora.

Zylbersztajn, A. (2015). *Cultura maker e educação* [Maker culture and education]. Senac São Paulo.

# "AI Talent for All": Sociotechnical Imaginaries in Korea's 2025 AI Education and Regulation Strategy

Seri Yoon, Waseda University, Japan

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

In November 2025, South Korea announced under the slogan "AI talent for all" an AI education package including a 5.5-year bachelor's-to-PhD fast-track, investment in "AI highways" data centers, an AI regulatory "rationalization" roadmap, and a lifelong AI education strategy. Ultimately, these measures form a national plan for Korea to become a "top three AI power" by 2030. The plan emphasizes innovation, competitiveness, and talent production. However, it remains unclear how education is reimagined or what this implies for equity and public accountability. Using Sheila Jasanoff's concept of sociotechnical imaginaries, this study examines how Korea's strategy aligns education, regulation, and infrastructure around AI competitiveness. Methodologically, the study uses qualitative critical policy analysis and discourse analysis on a purposive corpus of authoritative agenda-setting official documents from November to December 2025. The corpus includes high-level speeches, cross-ministerial roadmaps, and official press releases explicitly linked to the "AI talent for all" and "AI G3" initiative. Data are analyzed through a hybrid coding approach integrating deductive categories derived from Jasanoff's framework with inductive thematic coding. Analysis suggests education is cast as an AI workforce production system, while questions of educational inequality, labor conditions, and democratic oversight remain marginal. By conceptualizing Korea's strategy as a sociotechnical imaginary, the study offers a lens to analyze AI initiatives in other countries. The study argues AI education policy should be imagined not only as an instrument of competitiveness, but also as a space for critical digital literacy, institutional accountability, and inclusive futures resonating with UNESCO and OECD priorities.

*Keywords:* AI education policy, national AI strategy, AI governance, South Korea

iafor
The International Academic Forum
www.iafor.org

## Introduction

Artificial intelligence has become a central field through which governments imagine economic growth, industrial transformation, national security, and future social development. This shift has also changed the role of education within national technology policy. Education is expected to prepare individuals to use emerging technologies while developing the expertise and professional capacity required for national AI development. AI education is embedded within wider systems of infrastructure, industrial strategy, labor formation, data governance, and regulation. Critical scholarship similarly emphasizes that AI technologies operate within institutional, political, economic, and cultural processes rather than entering education as neutral technical tools (Williamson & Eynon, 2020).

South Korea offers a significant context for examining this relationship because its recent AI strategy explicitly links technological capacity with national competitiveness. In 2025, the government placed the ambition of becoming one of the world's top three AI powers at the center of a broader national transformation strategy encompassing the AI Highway, advanced computing capacity, domestic AI models, core AI talent, regulatory innovation, and the diffusion of AI across society (Ministry of Science and ICT, 2025b). Education became a particularly visible part of this project in November 2025, when the Ministry of Education announced its AI for All strategy. The strategy connects basic AI education across the life course with advanced talent cultivation, including a proposed five-and-a-half-year pathway from undergraduate study to doctoral training. AI education is organized along a continuum from general literacy to specialized expertise, linking inclusive access with accelerated talent production (Ministry of Education, 2025).

This relationship raises longstanding questions about the purposes of education during technological and economic change. AI literacy can support individual agency and participation in increasingly automated societies, while advanced education and research contribute to technological capacity. Yet access to digital devices, computing systems, and platforms does not guarantee equal participation. Differences in infrastructure, skills, and institutional support shape who benefits from technological change (OECD, 2023). UNESCO and the OECD similarly emphasize literacy, privacy, equity, stakeholder participation, labor rights, and institutional accountability as dimensions of responsible digital transformation (Miao & Holmes, 2023; OECD, 2023; UNESCO, 2021). These concerns become particularly relevant when education develops alongside the infrastructure and regulatory systems supporting national AI expansion.

The Korean strategy brings these dimensions together within a common national project. Computing capacity, data resources, and the National AI Computing Center are presented as foundations of technological transformation, while the AI Basic Act and its implementation framework seek to combine industrial promotion with safety and trust (Ministry of Science and ICT, 2025a, 2025b). The key question is how these objectives are related and which carry greater institutional weight. Equity may be recognized while investment remains concentrated on competitiveness; safety may be acknowledged while regulatory speed is prioritized; and AI education may be presented as a public benefit while assigning particular strategic value to advanced technological talent. A sociotechnical imaginary perspective makes it possible to examine the future that renders these investments necessary, the actors expected to build it, and the institutional arrangements through which it is constructed (Jasanoff & Kim, 2015).

Korea's late 2025 AI strategy can be read as a future-making project in which education, infrastructure, and governance are aligned around national AI competitiveness. The AI G3 objective identifies the desired national position, while the AI Highway and talent policies define the material and human capacities expected to sustain it. Educational institutions are positioned from general AI literacy to advanced expertise, and regulatory reform establishes conditions for technological development. This arrangement also raises questions about what becomes less visible within a competitiveness-centered vision, including unequal access across regions and social groups, changing labor conditions, and concerns over privacy, accountability, and public participation (OECD, 2023; UNESCO, 2021).

Existing research on AI in education has examined teaching, learning, competence, ethical risk, and digital governance, while critical scholarship has emphasized the need to situate AI within wider social and political processes (Williamson & Eynon, 2020). International frameworks further stress human-centered AI literacy, equitable infrastructure, and responsible governance (Miao & Holmes, 2023; Miao et al., 2024; OECD, 2023). This study extends these discussions by examining what happens when education becomes institutionally connected to national AI infrastructure and regulatory strategy under an explicit project of global technological competition.

The study asks how South Korea's 2025 AI education and regulation strategy constructs a sociotechnical imaginary in which education, infrastructure, and governance are reorganized around national AI competitiveness. It examines what this construction implies for the meaning of AI talent, the institutional role of education, and the place of equity and public accountability within the imagined future. Focusing on authoritative government documents associated with the AI for All and AI G3 initiatives, the analysis examines how policy language defines desirable futures, legitimate actors, governance priorities, and material infrastructures. The analysis demonstrates how national AI strategies shape the purposes assigned to education and the futures that public institutions are expected to create.

## Methodology

This study employs a qualitative research design combining critical policy analysis, discourse analysis, and qualitative document analysis to examine how South Korea's 2025 artificial intelligence strategy constructs a national vision of education, technological capacity, and governance. The analysis focuses on how authoritative policy texts define educational purposes, desirable forms of talent, relationships between education and technological development, and competing public values within a broader vision of an AI future. Policy documents are treated as substantive qualitative data because they reveal how institutions define problems, justify intervention, assign responsibilities, and establish priorities (Bowen, 2009).

The study examines a bounded policy period from November to December 2025, when AI education, talent cultivation, infrastructure, regulation, and national competitiveness became increasingly connected. During this period, the Ministry of Education announced the AI for All strategy, while the Ministry of Science and ICT advanced the implementation framework for the AI Basic Act and a broader national AI strategy (Ministry of Education, 2025; Ministry of Science and ICT, 2025a). The empirical corpus consists of 15 authoritative official documents selected purposively for their relevance to the research questions. These include government speeches, cross-ministry roadmaps, implementation materials, and press releases issued primarily by the two ministries and national AI governance bodies. The documents address the AI for All strategy, the AI G3 objective, and the broader policy architecture connecting

education, talent, infrastructure, and AI governance. The AI for All strategy serves as the central educational text, while the remaining materials situate it within Korea's wider AI policy framework. Media commentary, corporate promotional materials, and individual opinion pieces were excluded (Ministry of Education, 2025; Ministry of Science and ICT, 2025b).

The concept of sociotechnical imaginaries provides the main theoretical framework, directing attention to institutionally sustained visions of desirable futures that connect technological development with social order and political priorities (Jasanoff & Kim, 2015). The framework was operationalized through six deductive coding dimensions: future visions, legitimate actors, institutional arrangements, material infrastructure, governance rationales, and margins and absences within the dominant policy vision.

**Table 1**

*Analytical Dimensions of the Sociotechnical Imaginary*

| Dimension | Focus of Analysis |
|---|---|
| Future visions | Desired National & Social Future of AI |
| Legitimate actors | Government, Universities, Researchers, Industry, Teachers, Learners, Citizens |
| Institutional arrangements | Roles of Educational, Governmental, & Industrial Institutions |
| Material infrastructure | AI Highway, GPUs, data centers, computing resources |
| Governance rationales | Regulation, innovation, safety, trust, risk |
| Margins and absences | Equity, Accountability, Participation, Critical Literacy, etc. |

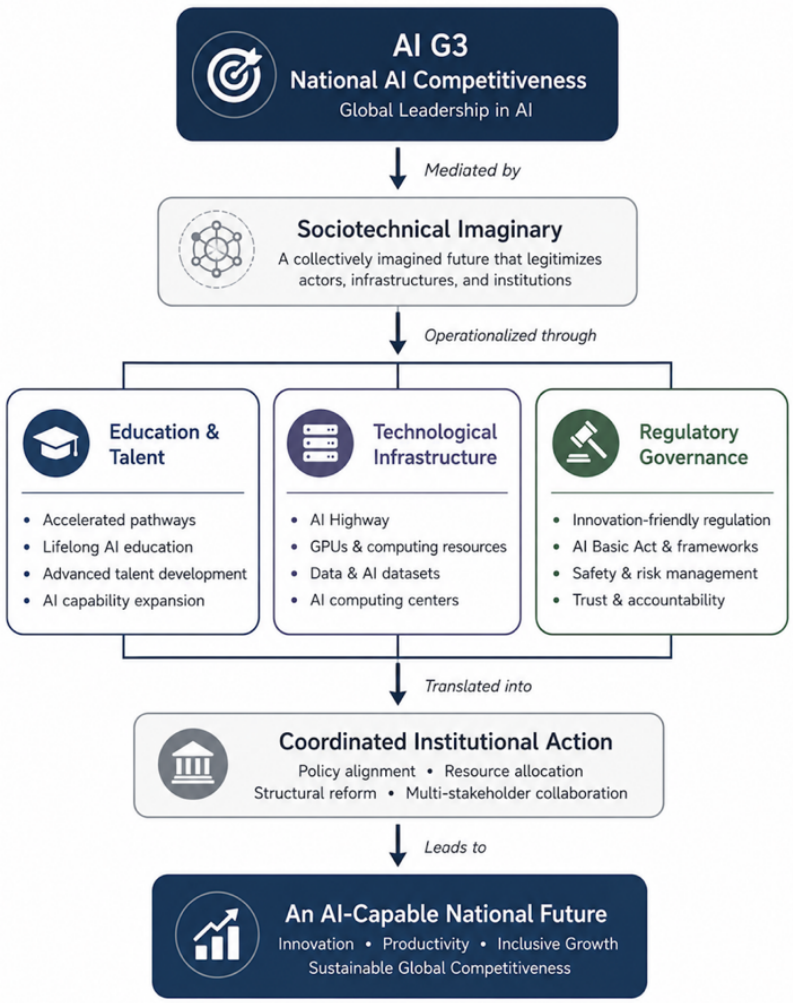
The analysis involved systematic reading and coding of the corpus to identify patterns across the six deductive dimensions of the sociotechnical imaginary. It also examined recurring themes concerning AI talent, education and workforce development, infrastructure, regulation and innovation, and the relative institutional weight given to equity and public accountability. Comparative analysis across documents and institutions identified common patterns while retaining meaningful differences. Particular attention was given to distinguishing between the presence of a policy value and its institutional weight, assessing whether values such as inclusion, safety, ethics, and equity were reflected in policy objectives, institutional responsibilities, or implementation arrangements. Korean originals served as primary sources where available, with official English versions used to cross-check terminology. This approach follows critical policy analysis by examining how policy discourse organizes priorities and legitimizes particular understandings of problems and solutions (O'Connor & Rudolph, 2023).

The study focuses on policy discourse rather than the long-term effectiveness of Korea's AI education strategy or its causal effects on educational inequality, labor market outcomes, and regulation. It examines the underlying logic through which education, talent, infrastructure, and regulation are integrated into a common sociotechnical imaginary.

## Results and Findings

The analysis reveals that the Korean AI policy imaginary is organized around national competition, technological acceleration, human resource development, and institutional transformation. Across the documentary corpus, education is positioned within a broader national architecture linking talent, computing capacity, industrial transformation, and regulatory reform to Korea's ambition to become one of the world's leading AI powers. The AI G3 vision establishes the desired national position, while education and infrastructure provide the human and material capacity through which that future is expected to become achievable (Jasanoff & Kim, 2015; Ministry of Education, 2025; Ministry of Science and ICT, 2025b). Within this configuration, AI talent is constructed as a strategic national asset. The Ministry of Education presents intensifying global AI competition as a condition in which delay may produce generational disadvantage and connects talent cultivation to national survival. Students and researchers are positioned as individuals pursuing knowledge and as resources expected to strengthen the technological position of the nation (Ministry of Education, 2025).

**Figure 1**
*From National AI Vision to Institutional Action*

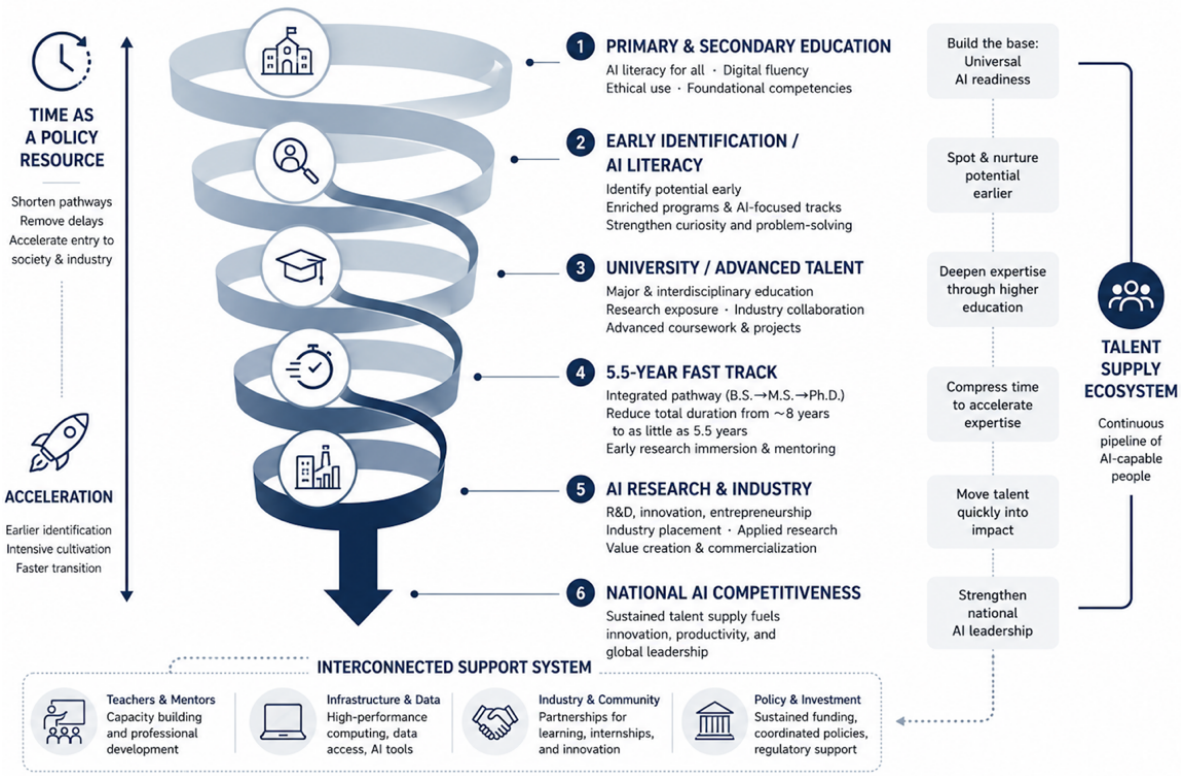



The connection between education and the AI G3 objective reinforces this construction of talent. The Ministry of Education presents talent development alongside infrastructure and data expansion as necessary to national AI advancement. Education is incorporated into a broader conception of national capacity in which human and technological capabilities develop together

(Ministry of Education, 2025; Ministry of Science and ICT, 2025b). Policy measures include earlier identification of students with AI capabilities, specialized education, support for advanced research, and an integrated pathway that could reduce the period from undergraduate entry to doctoral completion from approximately eight years to as little as five and a half years. Financial support for selected science and engineering students and expanded institutional support for research careers further reinforce this model (Ministry of Education, 2025).

**Figure 2**

*The Accelerated and Interconnected AI Talent Production System*




The analytical significance of these measures lies in their temporal logic. Time becomes a policy resource: the issue is not only whether advanced talent can be produced, but how rapidly it can enter research and professional roles. The shortened doctoral pathway most clearly expresses this orientation, linking educational acceleration to earlier participation in national technological development (Ministry of Education, 2025).

Education is also organized as an interconnected system for producing and renewing AI-capable labor. The policy extends from basic literacy in schools through higher education, specialist training, workplace learning, and lifelong education. At the university level, pathways range from AI education for non-majors to regional talent hubs and advanced research routes, while university-industry agreements, company-based graduate education, internships, and vocational restructuring bring education closer to employment (Ministry of Education, 2025). The learner is situated within a continuous pathway linking education, research, and productive work.

The findings do not suggest that workforce preparation is an inappropriate purpose of education. Rather, the documentary analysis reveals the growing dominance of a production-oriented logic. Education becomes increasingly legible to the state through what it can supply, while the value of institutions is linked to their capacity to produce people who contribute to an AI-

transformed economy. Broader purposes remain visible in references to AI ethics, critical thinking, and regional inclusion, but they do not organize the strategy to the same extent as talent cultivation and competitiveness (Ministry of Education, 2025).

Technological infrastructure gives this imaginary a material form. The AI future is tied to computing, semiconductors, and data centers rather than presented only through aspirations concerning innovation and competitiveness. The AI Highway frames computing capacity as foundational national infrastructure, with a stated objective of securing more than 50,000 graphics processing units through public-private collaboration and a National AI Computing Center expected to contribute more than 15,000 advanced graphics processing units by 2028. These commitments turn the AI G3 ambition into a material project requiring substantial investment in computing capacity and related infrastructure (Ministry of Science and ICT, 2025b).

The relationship between infrastructure and education is particularly revealing. Infrastructure and data expansion are presented alongside talent cultivation, while regional universities are expected to connect AI education with computing resources and local industry. The policy imaginary links deficits in computing, expertise, and regional capacity to measures including the AI Highway, accelerated educational pathways, and regional education-industry networks. Material and social arrangements are developed together around the same national objective (Jasanoff & Kim, 2015; Ministry of Education, 2025; Ministry of Science and ICT, 2025b). The state remains central to this process. Although the strategy relies on public-private cooperation, public institutions establish direction and capacity, while private actors contribute to implementation and commercialization (Ministry of Science and ICT, 2025b).

Regulatory transformation represents another dimension of the project. The AI Basic Act and its implementation framework retain provisions concerning fundamental rights, safety, and high-impact AI, including algorithmic impact assessments, while emphasizing flexibility, reduced uncertainty, and limits on unnecessary burdens for companies. Regulation is framed as part of the institutional infrastructure for technological development rather than simply as an external constraint. Rights and safety remain on the policy agenda, but within a political temporality shaped by technological urgency and national competitiveness (Ministry of Science and ICT, 2025a, 2025b).

Questions of equity, labor conditions, democratic oversight, and public accountability remain comparatively peripheral within this dominant imaginary. These concerns are not absent: the Ministry of Education addresses regional disparities, lifelong learning, and AI ethics, while national AI governance includes accessibility, social inclusion, and human-centered AI (Ministry of Education, 2025; Ministry of Science and ICT, 2025b). The finding concerns hierarchy rather than absence. Talent, infrastructure, industrial transformation, and competitiveness are translated into institutional pathways, measurable targets, dedicated organizations, and material investments. Democratic participation, the distribution of authority over educational transformation, and labor conditions receive less extensive institutionalization.

Educational inequality illustrates this distinction. The strategy recognizes the concentration of AI infrastructure and educational opportunity in the capital region and proposes regional centers and university hubs. Yet the dominant response remains the expansion and redistribution of AI capacity. Less attention is given to whether different social groups have equal influence over the purposes of AI education, experience similar educational pressures, or face new hierarchies through competition for advanced talent. Equity is primarily constructed

as access to participation in the AI transformation rather than as a reconsideration of how that transformation should be governed (Ministry of Education, 2025).

A similar pattern appears in labor policy. The strategy addresses workforce demand through skills development, practical experience, company involvement in training, and accelerated movement into professional roles. Employment security, professional autonomy, and working conditions receive less attention. The policy imaginary is consequently more developed around workforce supply than around the social conditions of future AI-related work (Ministry of Education, 2025).

Public accountability occupies a similarly complex position. The regulatory framework includes consultation with industry, academia, and civil society, and the National AI Strategy Committee addresses social, safety, and public-sector concerns. Yet the dominant institutional language emphasizes coordinated execution, rapid implementation, national competitiveness, and measurable outcomes. The analyzed documents provide less detail on how educators, workers, and affected communities can exercise sustained influence over the direction of national AI transformation (Ministry of Science and ICT, 2025a, 2025b).

Critical digital literacy is also significant. The Ministry of Education includes ethical AI use, responsible engagement, humanities education, and critical thinking, preventing AI for All from being reduced to technical skills acquisition. At the same time, the institutional architecture surrounding advanced talent is more extensive and more directly connected to national economic objectives. Literacy functions primarily as a supporting educational value, while talent production operates as a central strategic mechanism (Ministry of Education, 2025). This relative emphasis distinguishes the Korean strategy from international frameworks that place greater weight on human-centered, ethical, technical, and civic dimensions of AI competence and on equity and institutional accountability in digital education (Miao et al., 2024; OECD, 2023).

The five findings converge around a common structure: AI talent is constructed as national capacity, education as a continuous and increasingly accelerated system for developing that capacity, infrastructure as its material foundation, and regulation as a mechanism for securing trust while remaining compatible with rapid innovation. Equity, public accountability, labor conditions, and critical literacy remain present but carry less institutional weight.

The AI G3 objective functions as an organizing vision that connects policies that might otherwise appear separate. Shortened doctoral pathways, regional university investment, computing centers, and regulatory reform become parts of the same national project because each contributes to a shared vision of technological leadership. The sociotechnical imaginary coordinates social and technical change by defining what the future requires and assigning institutions responsibility for producing it (Jasanoff & Kim, 2015; Ministry of Education, 2025; Ministry of Science and ICT, 2025b).

The phrase AI for All contains a related tension. The strategy expands access through lifelong learning, regional provision, and opportunities for non-majors, yet participation remains closely connected to technological competitiveness. Citizens are primarily imagined as competent users, adaptable workers, researchers, professionals, and producers of economic and technological value (Ministry of Education, 2025). Inclusion can thus operate within a competitive sociotechnical imaginary without altering its central logic. The key issue is

whether values such as equity, participation, and accountability acquire sufficient institutional weight to shape the direction of technological transformation.

The findings show how Korea's 2025 AI strategy connects domains traditionally governed separately. Education becomes part of technological strategy, infrastructure part of national economic transformation, and regulation part of innovation policy. Universities serve educational, research, regional, and talent-development functions, while citizens enter a lifelong AI learning environment and selected individuals are positioned as strategic experts. These relationships create an institutional architecture through which the imagined AI future can be actively constructed (Jasanoff & Kim, 2015).

The central finding is that education has become a principal mechanism through which Korea's AI-centered future is made socially possible. The policy assigns educational institutions responsibility for developing the people, competencies, research capacity, and forms of participation required for technological transformation. The unresolved question is how broadly education can be defined when national technological competition becomes one of its most powerful organizing forces.

## Implications and Conclusion

The findings of this study extend beyond Korea's 2025 AI strategy, showing that AI education policy is inseparable from the political and economic future it helps shape. When education, infrastructure, regulation, and industrial development are coordinated through a common national vision, education policy also shapes the society that technological transformation produces. The central challenge is to develop technological capability without reducing education to national competition. Sociotechnical imaginaries are useful here because technological futures reflect assumptions about desirable institutions, legitimate actors, and collective purposes (Jasanoff & Kim, 2015). While only a small proportion of workers may require advanced AI-specific capabilities, most will need basic digital competence, critical problem-solving skills, and the ability to work with AI systems (OECD, 2026). Given persistent differences in digital experience and access to workplace AI training across Korean firms (OECD & Korea Labor Institute, 2025), lifelong AI education must address not only skills but also the resources and institutional support required for meaningful participation.

Educational equity remains a central concern. The benefits of AI education may depend on unequal access to infrastructure, research capacity, financial resources, and digital experience. This is particularly relevant to Korea's regional strategy: expanding AI education beyond the capital region does not ensure equality without sustained computing resources, qualified educators, research networks, and pathways to professional opportunities (OECD, 2023). The same concern applies to accelerated talent production. The five-and-a-half-year undergraduate-to-doctoral pathway may respond to international AI competition and benefit exceptionally prepared students, but speed should not become a universal measure of educational effectiveness. Advanced education also requires disciplinary depth, experimentation, failure, and ethical reflection. As universities are increasingly expected to educate students, produce researchers, and collaborate with industry, technological capacity should not come at the expense of fundamental research, critical inquiry, or institutional freedom to question technological development.

The AI Highway should be understood as more than a system of computing resources. Graphics processing units, data centers, cloud resources, and research facilities acquire public value

through the institutions and people able to use them. Investment in physical infrastructure needs to be accompanied by investment in educators, researchers, public institutions, and regional networks. Digital infrastructure is not educationally neutral; its effects depend on how access is distributed and who has the capacity to determine its uses. The success of national AI strategy should be assessed through more than computing capacity, numbers of trained specialists, international rankings, or industrial investment. Evaluation should also consider who gains access to high-quality AI education, whether regional and socioeconomic gaps are narrowing, and whether teachers, learners, and institutions retain meaningful agency and public accountability (OECD, 2023).

Trust, regulation, and participation raise related questions. Public protection should not be treated simply as a cost of technological progress, as accountability and oversight can strengthen both legitimacy and institutional certainty (UNESCO, 2021). Korea's regulatory framework need not choose between flexibility and precaution; the more important question is whether innovation speed becomes the primary measure of effective governance. The current policy imaginary also emphasizes workforce supply more than participation in technological governance, despite the knowledge that students, teachers, workers, and communities can contribute to understanding AI's social effects (Doellgast et al., 2025).

National competitiveness should remain one legitimate public objective, not the unquestioned measure through which all other objectives are organized. Countries have understandable reasons to invest in AI, but competitiveness can become the dominant imaginary through which educational value, regulatory quality, infrastructure, and social participation are interpreted. A broader framework should evaluate technological development through capability, equity, rights, democratic legitimacy, and long-term social value. Within this framework, AI for All acquires a more substantive meaning. Universal AI education requires more than expanding access to basic knowledge or increasing participation in AI education. Advanced opportunities should not be determined primarily by geography or social advantage, and specialized expertise should develop alongside broader human and social capabilities. Innovation should remain open to scrutiny, while public investment in computing capacity should generate widely distributed benefits. Inclusion, in this sense, concerns the meaningful choices, capabilities, and forms of participation available to people as technological transformation proceeds.

The implications extend beyond South Korea. Many governments are developing national AI strategies that combine education, talent, infrastructure, industrial policy, and regulation. Although institutional arrangements differ, the questions raised here remain relevant: What kind of future is presented as desirable, who is authorized to construct it, and which public values receive institutional support rather than remaining primarily rhetorical? Sociotechnical imaginaries can illuminate not only national visions but also the distribution of authority and possibility within them.

The study makes a theoretical contribution by extending sociotechnical imaginaries more explicitly into AI education policy. Education is often treated as a response to technological change, yet the findings show that it can also be one of the institutions through which an imagined technological future is produced. Schools create basic forms of technological participation, universities develop expertise and research, lifelong education supports adaptation, and talent policies identify the actors expected to occupy strategically important positions. Education does not simply prepare society for a technological future that exists independently; it helps determine which futures become institutionally possible. The Korean case demonstrates why educational policy, material infrastructure, and governance should be

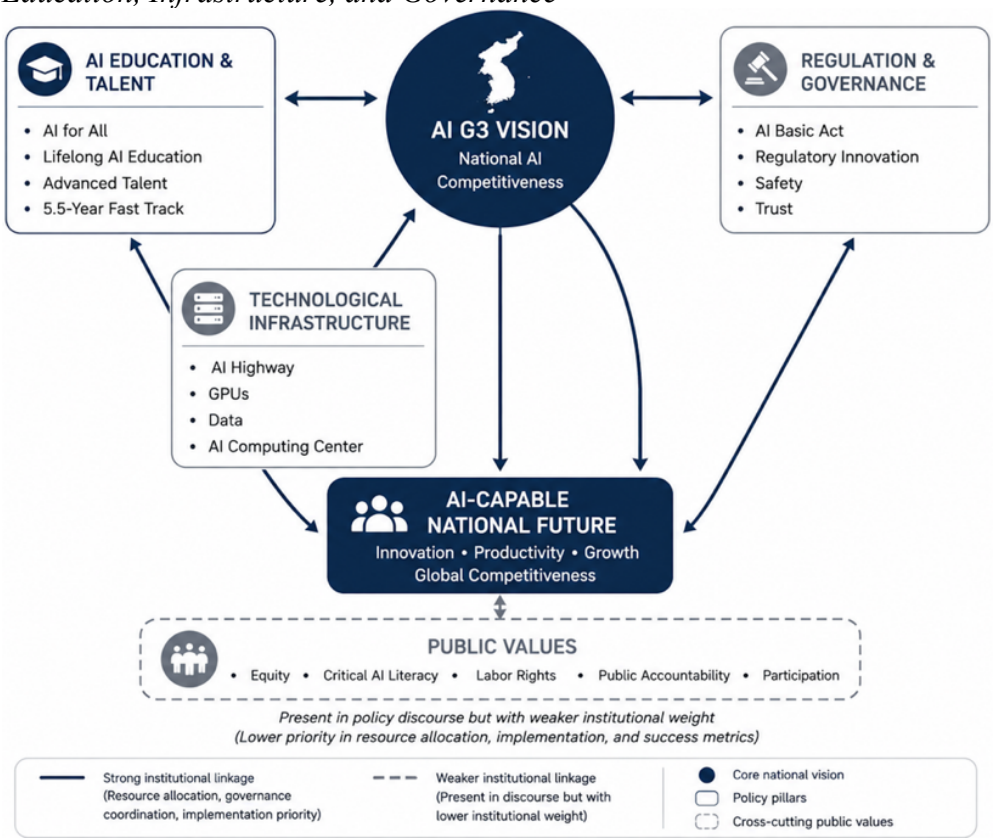
analyzed together. Computing investments provide the technical foundation for national ambitions, educational institutions produce the knowledge and expertise needed to use that infrastructure, and regulatory arrangements establish the conditions under which technological development proceeds. A sociotechnical imaginary gains practical force when future narratives become connected to systems of investment, education, and governance.

The methodological contribution lies in distinguishing the presence of a policy value from its institutional weight. Critical analysis should neither assume that a value is absent because it does not dominate a policy document nor treat the appearance of terms such as inclusion, ethics, safety, or fairness as evidence that they determine policy direction. The more important question is whether such values shape resource allocation, implementation structures, accountability mechanisms, and definitions of policy success. This distinction provides a useful approach to policies that pursue economic, technological, educational, and social goals simultaneously.

South Korea's 2025 AI education and regulation strategy can be understood as a sociotechnical imaginary through which a particular national future is made desirable and institutionally actionable. The analysis found that AI talent is constructed as a strategic national asset, while education is increasingly organized as a continuous system for developing and accelerating AI capability. Technological infrastructure provides the material foundation, and regulatory reform seeks to create conditions compatible with rapid innovation. Equity, labor conditions, critical literacy, and public accountability remain present within the policy environment, but they carry less organizing power than talent, infrastructure, acceleration, and competitiveness.

**Figure 3**

*South Korea's 2025 AI Sociotechnical Imaginary: From National Competitiveness to Education, Infrastructure, and Governance*



These findings do not suggest that Korea's AI strategy is merely a project of workforce production. Its ambition lies in coordinating AI education, technological capacity, advanced expertise, and regional development at a national scale. The critical question is how this capacity will be governed and for what purposes. Education should serve more than the needs of a technologically competitive state. AI education should help people participate in technological change while also understanding it, exercising judgment, and shaping the institutions that govern it. A society is not truly AI capable simply because it has advanced models, computing infrastructure, and specialized workers. It must also have citizens and institutions able to question how technology distributes power, opportunity, responsibility, and risk.

The phrase AI for All contains a possibility that is broader than the current logic of talent production. It can describe a future in which everyone is expected to adapt to AI, or it can describe a future in which everyone is given a meaningful capacity to participate in shaping it. The distinction between these two futures is ultimately a question of education and governance. The first requires skills. The second requires skills together with critical literacy, institutional accountability, public participation, and equitable access to the resources through which technological futures are created. The long-term public value of Korea's AI strategy will depend not only on whether it succeeds in producing competitive AI talent but also on whether the future constructed through that talent remains open to the people who are expected to live within it.

**Table 2**

*Divergent Interpretations of "AI for All"*

| Dimension | AI for All as Talent Production | AI for All as Democratic Participation |
|---|---|---|
| Primary purpose | Workforce preparation | Participation in shaping technological futures |
| Education | Skills and adaptation | Critical literacy and agency |
| Social role | Economic participation | Civic and institutional participation |
| Governance | Innovation and competitiveness | Accountability and public oversight |
| Public values | Secondary / weaker institutional weight | Embedded in institutional decision-making |
| Desired outcome | AI-capable workforce | AI-capable and AI-governing society |

## References

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Doellgast, V., Appalla, S., Ginzburg, D., Kim, J., & Thian, W. L. (2025). *Global case studies of social dialogue on AI and algorithmic management*. International Labour Organization. https://doi.org/10.54394/VOQE4924

Jasanoff, S., & Kim, S. H. (Eds.). (2015). *Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226276663.001.0001

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Miao, F., Shiohira, K., & Lao, N. (2024). *AI competency framework for students*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000391105

Ministry of Education. (2025, November 10). *AI for all: Strategy for cultivating artificial intelligence talent*. https://english.moe.go.kr/boardCnts/viewRenewal.do?m=0202&s=english&page=1&boardID=254&boardSeq=104821&lev=0&opType=N

Ministry of Science and ICT. (2025a, November 12). *MSIT announces legislative notice for the enforcement decree of the AI Basic Act to foster the AI industry and build a foundation for safety and trust*. https://www.msit.go.kr/eng/bbs/view.do;jsessionid=aBO-HMW2y4oNRpiPJjDbV_NuiV_0iPtKthgXL-ul.AP_msit_1?sCode=eng&mPid=2&mId=4&bbsSeqNo=42&nttSeqNo=1191

Ministry of Science and ICT. (2025b, September 8). *National AI Strategy Committee launched as Korea's top level AI policy body*. https://msit.go.kr/eng/bbs/view.do?sCode=eng&mPid=2&mId=4&bbsSeqNo=42&nttSeqNo=1165

O'Connor, K., & Rudolph, S. (2023). Critical policy analysis in education. In K. Hytten (Ed.), *Oxford Research Encyclopedia of Education*. https://doi.org/10.1093/acrefore/9780190264093.013.1831

OECD. (2023). OECD digital education outlook 2023: *Towards an effective digital education ecosystem*. OECD Publishing. https://doi.org/10.1787/c74f03de-en

OECD. (2026). *AI and skills: What we know so far*. OECD Publishing. https://doi.org/10.1787/f843b352-en

OECD, & Korea Labor Institute. (2025). *Artificial intelligence and the labour market in Korea*. OECD Publishing. https://doi.org/10.1787/68ab1a5a-en

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology, 45*(3), 223–235.

# Surpass Chatbots and Human Teachers? Explore the Effectiveness of Conversational AI in Language Learning

Yitian Sun, The Hong Kong Polytechnic University, Hong Kong, China
Shimeng Wu, The Hong Kong Polytechnic University, Hong Kong, China
Cong Qi, The Hong Kong Polytechnic University, Hong Kong, China

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Conversational Artificial Intelligence (ConvAI) has evolved from early rule-based systems and simple chatbots (e.g., Siri) to complex learning tools enabled by large language models, opening up new possibilities for language education. ConvAI is a type of AI that enables computers to understand, process, and respond to human language in a natural and human-like way. The introduction of ConvAI to language learning is especially timely and meaningful, as it overcomes the limitations of traditional learning methods, such as limited practical scenarios, mechanical or unengaging feedback, and high human costs. ConvAI systems, sometimes represented as avatar, can simulate real-life conversations, provide immediate correction, and offer adaptive learning options as well as context-aware feedback. This study evaluated the effectiveness of ConvAI in language education from two ConvAI-enabled language learning tools: Loora and Talkpal. Detailed differences between these two platforms are summarized. Moreover, the overall analysis reveals that ConvAI-enabled learning tools can break temporal and spatial barriers for language learning, offer highly personalized and adaptive learning experiences, provide vivid, immediate, and granular feedback to language learners, democratize learning costs, and ultimately enhance students' grammatical accuracy and overall English proficiency. Our study also highlights key challenges in ConvAI-enabled language learning tools, such as limited depth, lack of human interaction, difficulties in fully replicating cultural and contextual nuances, and data privacy risks. Finally, we compare learning practices between human teachers and ConvAI avatar teachers. The comparison results arouse educators' attention on effective Human-ConvAI collaboration in language learning.

*Keywords:* conversational AI, language learning, effectiveness

iafor
The International Academic Forum
www.iafor.org

## Introduction

Conversational Artificial Intelligence (ConvAI) enables computers to understand, process, and respond to human language naturally through text or speech. It integrates multiple technologies including Natural Language Processing (NLP) with Automatic Speech Recognition (ASR), Machine Learning, Deep Learning and advanced image processing (Zawacki-Richter et al., 2019), with an aim to achieve seamless interaction between humans and machines (Chigbu & Makapela, 2025). ConvAI's evolution has progressed from early rule-based systems like ELIZA (Weizenbaum, 1966) to contemporary large language model-driven systems (e.g., GPT series), which possess improved comprehension, contextual memory, and generative capabilities (Wu et al., 2024). ConvAI is increasingly moving from general purpose chatbots to specialized applications, and language education field is emerging as a key application domain.

The integration of ConvAI into language education offers significant advantages over traditional approaches. It addresses global teacher shortages, provides timely feedback that guarantee learning effectiveness, and delivers scalable, personalized instruction (Alkishri et al., 2025). Moreover, ConvAI makes language learning more resourceful, more interactive, thus reducing language learning anxiety (Ji et al., 2022). Language learning is about continuous interaction with teacher, Interaction Theory (Kinginger, 2001) well explains the roles of language teachers as well as ConvAI tools for language learning. Though ConvAI has significant benefits for language learners and teachers, like other AI tools, it also carries risks and challenges. Current ConvAI technology exhibits limitations in handling specialized subjects and cultural nuances. ConvAI models may also suffer from hallucinations and algorithmic bias, demonstrate emotional deficits compared to human tutors in expressing sentiment, and raise ethical concerns regarding user data privacy (Aslan et al., 2025).

To evaluate the effectiveness and limitations of ConvAI in language education, we selected two of the most typical ConvAI based learning tools in the market: Loora and Talkpal and have a thorough investigation and comparison between these two. The functionalities and impact on users are compared and discussed. At the same time, we also summarized the major challenges and risks in using these ConvAI tools. Furthermore, this paper compares the differences between ConvAI and human tutors, exploring potential avenues for complementary strengths. Our research results can provide a reference for educators, technology developers, and relevant stakeholders regarding the appropriate use of ConvAI, thereby advancing the optimization and innovation of ConvAI in language education.

## Literature Review

### Existing Research on Conversational AI in Language Learning

Existing research indicates that ConvAI tools can create learning environments and opportunities for language learners that traditional teaching methods struggle to provide.

Recent studies have explored the integration of AI-supported communication tools in language learning. Shadiev et al. (2025) reviewed 31 studies published between 2010 and 2024 and identified a range of AI technologies used to support language learning. This includes conversational agents, speech-recognition systems, and advanced AI-powered chatbots. Their findings prove that these ConvAI tools have positive effects on speaking performance, communication competence, and learner motivation, while also highlight challenges such as

technical limitations and insufficient teacher support. Similarly, Dai and Wu (2025) found that AI-assisted speaking practice enhanced learners' engagement and oral performance, with real-time feedback emerging as a particularly valuable feature. What is more, AI-supported speaking tools are also associated with gains in language learning fluency, pronunciation, confidence, motivation, and willingness to communicate (Xing & Saeed, 2025). However, existing reviews also show that chatbot-supported language learning remains limited by speech-recognition errors, inconsistent corrective feedback, uneven dialogue quality, learner-level mismatch, and insufficient integration with teacher-guided pedagogy (Chua & Annamalai, 2025). Therefore, while current evidence supports the pedagogical potential of ConvAI, more practice-oriented evaluation is needed to determine how commercial platforms function in authentic learning contexts and whether they can provide sustained, culturally sensitive, and pedagogically reliable support alongside human teachers.

## Major Theories Suitable to Explain the Use ConvAI in Language Learning

Based on the investigation of past literature, we summarized four major theories suitable to explain the effectiveness of ConvAI in language learning. They are Task-Based Language Teaching, Willingness to Communicate Theory, Self-Determination Theory, and the Affective Filter Hypothesis. First, Task-Based Language Teaching (TBLT) focuses on the use of authentic language to complete meaningful tasks in the target language, such as visiting a doctor or calling customer service for help. Assessment is primarily based on task outcomes rather than on accuracy of language use (Ellis, 2022). ConvAI is believed to be able to provide interactive real-world speaking practices for the language learners, to cope with scenario and task-based language learning. Second, in the second language learning context, MacIntyre et al. (1998) cited Willingness to Communicate (WTC) Theory to explain the extent to which learners are prepared to communicate with a specific interlocutor in a second language at a given moment. The willingness is influenced by situational WTC such as immediate states, self-confidence, and social/environmental influences (Zhang et al., 2018). WTCT makes a key concept to understand spoken participation in a second language. ConvAI tools using robot or humanoid robot as significant interlocutor reduces the pressure from peer, group and classroom environment, thus boost confidence and minimize anxiety when learning secondary language. Third, Deci and Ryan (1985) introduced Self-Determination Theory, which focuses on understanding human's motivation and personality. They believe human's motivation ranges from unmotivated to fully self-determined, while intrinsic motivation represents the highest form of self-determination. Intrinsic motivation is embedded in enjoyment, challenge, or inherent satisfaction of the activity itself. If a learner intrinsically enjoyed "playing" around the ConvAI tool, satisfied with the personalized and immediate feedback, then they will be greatly motivated to learn better. Last, the Affective Filter Hypothesis is a concept in second language acquisition developed by linguist Stephen Krashen. It proposes that a learner's emotional states, such as anxiety, fear, embarrassment, motivation, self-confidence, act as an invisible psychological "filter" that can either block or allow the language input necessary for fluency (Krashen, 1986). A low psychological filter allows more input while a high filter may hinder (Du, 2009). ConvAI platforms can effectively reduce emotional filtering and alleviate communication anxiety and the fear of negative evaluation by providing a private and non-judgmental practice environment. Therefore, the Affective Filter Hypothesis can also apply to the research context of ConvAI in language learning.

## Research Gaps and Our Research Interests

Existing literature demonstrated ConvAI's vast potential in language learning. However, there are also gaps in the past studies that our research would like to fill in.

First, there is a lack of systematic evaluation of commercialized ConvAI learning tools. Existing studies mostly use self-built experimental systems or single chatbots to evaluate their effectiveness. The comparison between several mature commercial platforms in language learning is rare. Second, research on human-AI comparison is insufficient. Although many studies have verified the effectiveness of using ConvAI tools, the study of directly comparing it with human teachers is limited. Language learning involves multi-dimensional elements such as emotional connection, cultural understanding and interpersonal interaction. This is not to be accomplished without human-AI collaboration. A systematic comparison of the differences between the two is important to understand the human-AI collaborative teaching model. Third, our study is among the pioneer studies to summarize the major theories used in ConvAI-based language learning context.

Based on the above discussions of the gaps in literature, this study selects two commercial ConvAI tools, Loora and Talkpal, and systematically evaluate their effectiveness and limitations in language learning. We further compare the differences between ConvAI tools and human teachers. We believe this study can provide some theoretical basis and practical summary for human-AI collaborative language teaching and learning.

# Case Study on Loora and Talkpal

## Case 1: Loora

Loora is an AI-powered English language learning platform that utilizes customized Large Language Models (LLMs) to simulate natural dialogue and support learners in a low-stakes, adaptive environment. Unlike traditional language learning applications that rely on pre-scripted content, Loora is an example of ConvAI applied to language education. It enables interactive practice and real-time feedback that were once limited to human-led instructions. Positioned as a premium speaking practice tool for serious English learners, Loora targets users with some language foundation who seek fluency for professional or academic purposes (Lunden, 2023). So far, Loora occupies the premium segment of the AI language learning market, distinguishing itself through specialized LLM customization and comprehensive feedback mechanisms.

### *Core Features and Pedagogical Design*

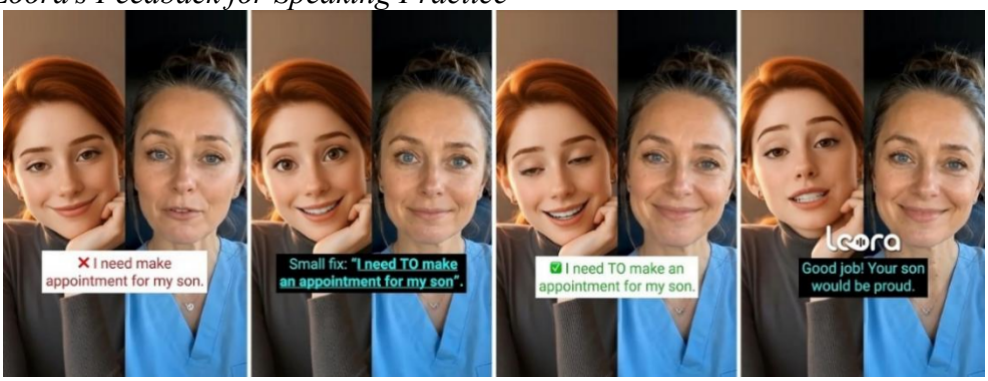
Loora's pedagogical design is built around three core technological features:

**Feature 1: Customized Large Language Models (LLMs) for Natural Dialogue.** Loora uses customized LLMs that have been specifically optimized for English language education. It developed its own LLMs in 2020 but also used a variety of third-party LLMs for different pedagogical tasks, such as initiating conversations, managing ongoing dialogues, tackling diverse topics, and assisting users with memory recall (Lunden, 2023). Customization was embodied when answers from the LLMs are contextually relevant, and able to adjust based on the current conversational status. For example, you can ask Loora's AI tutor to help you practice interview skills based on your job applied and with your current knowledge and skill

sets. The advantage of this is that users can practice English in real-life conversation with no judgement and pressure.

**Feature 2: Real-Time Feedback.** By combining ASR and NLP, Loora delivers immediate, actionable feedback. ASR can accurately transcribe a user's speech into text, which demonstrates consistent robustness even when faced with various accents (Maricheva & Kyrychenko, 2026). Then, NLP analyzes the transcribed text across several dimensions, such as grammatical errors, lexical appropriateness, and pragmatic competence (Stryker & Holdsworth, 2026). Loora provides pronunciation guidance by analyzing speech patterns against native speaker models and providing specific feedback on intonation and pronunciation. This real-time feedback is critical for preventing language from becoming "fossilized". Figure 1 illustrates a classical example of Loora's feedback mechanism. It shows how the system transforms a learner's spoken or written output into immediate corrective guidance. Such feedback enables learners to clearly recognize their linguistic errors and supports them in making timely self-corrections in speaking practice.

**Figure 1**
*Loora's Feedback for Speaking Practice*




**Feature 3: Contextual Understanding and Progress Visualization.** By learning contextual information from past conversational records, Loora can remember conversations and response consistently with previous interactions. This guarantees coherent and personalized dialogue (Alieva, 2025). The understanding of conversational context transforms Loora from a simple error-correction tool into a personalized AI tutor that "remembers" individual learner's preferences, challenges, and progress. What is more, the platform provides a comprehensive visual dashboard that automatically tracks and displays learning progress. This includes the number of days spent studying, status of completed lessons, and a line chart showing the number of lessons completed over time (Loora, 2023). Through real-time access to the learning history, learners can develop metacognitive awareness: understanding learning patterns, identifying areas that require additional practice, and setting evidence-based goals.

### *Limitations and Challenges*

Despite its many advantages, Loora faces some limitations. First, Loora is less effective for advanced or specialized conversational practice. Although Loora is positioned as a personal AI tutor capable of providing personalized conversations and real-time feedback, an evaluation by PracticeMe (2026) pointed out that its everyday conversation mode may not be deep enough for advanced learners. Secondly, given its functional limitations, Loora's high price (approximately US$119 per year) is difficult to justify, particularly when compared to human tutors, free AI alternatives such as ChatGPT, and rival AI tools like Talkpal. This price pressure

may lead to user churn, especially among price-sensitive learners. Thirdly, Loora may struggle to maintain conversational coherence in longer interactions or across different sessions. According to PracticeMe (2026), Loora sometimes forgets the content of previous discussions or repeat prompts across different sessions. Such limitations may disrupt the natural flow of conversation and reduce thematic consistency over time. Finally, user privacy risks are a major concern, as Loora collects a significant amount of sensitive data (e.g., voice, location, and search history) from learners.

**Case 2: Talkpal**

Talkpal is an AI-powered language learning platform offering multilingual support, flexible interaction modes, and a relatively affordable price. Unlike Loora, which focuses specifically on English speaking practice, Talkpal supports more than 57 languages, which include English, Spanish, French, German, and Chinese, and less commonly taught languages such as Arabic, Turkish, Persian, and Afrikaans. The platform building on ConvAI is based on GPT architecture, and is designed to simulate natural conversations, providing learners with interactive opportunities to practice speaking, writing, listening and reading in a low-pressure digital environment. Consequently, its market positioning differs from that of Loora: whilst Loora focuses on professional English tuition and in-depth feedback tailored to a single target language, Talkpal places greater emphasis on breadth, ease of use and a flexible, multilingual learning experience.

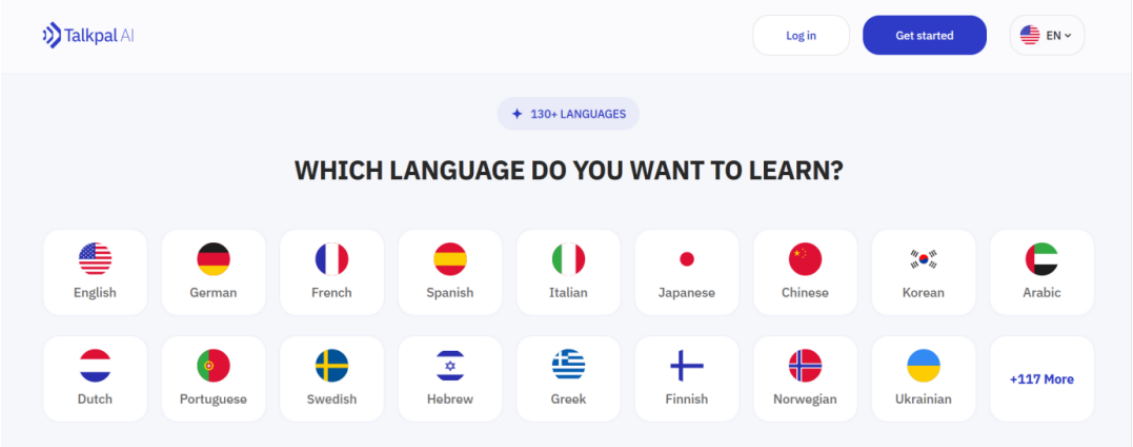
Talkpal's core value lies in providing learners with a more cost-effective and flexible alternative to traditional language tuition. According to Dunham (2026), the platform offers a variety of interactive modes, such as casual chat, role-play, character-based dialogue, debates, telephone conversations, sentence exercises, vocabulary exercises, conversation practice and picture description. These modes enable learners to practice the language in a variety of communicative contexts, rather than relying solely on vocabulary or grammar exercises. Therefore, Talkpal holds appeal for learners who wish to be involved with multiple languages, undertake flexible speaking practice, and utilize AI-based conversation tools at a lower cost.

***Core Features and Pedagogical Design***

Talkpal combines GPT-powered ConvAI into a multi-modal language learning structure designed to support real-time, immersive language practice. Through text and voice interactions, Talkpal enables learners to express themselves, receive feedback, and adjust their expressions with a continuous cycle of conversation.

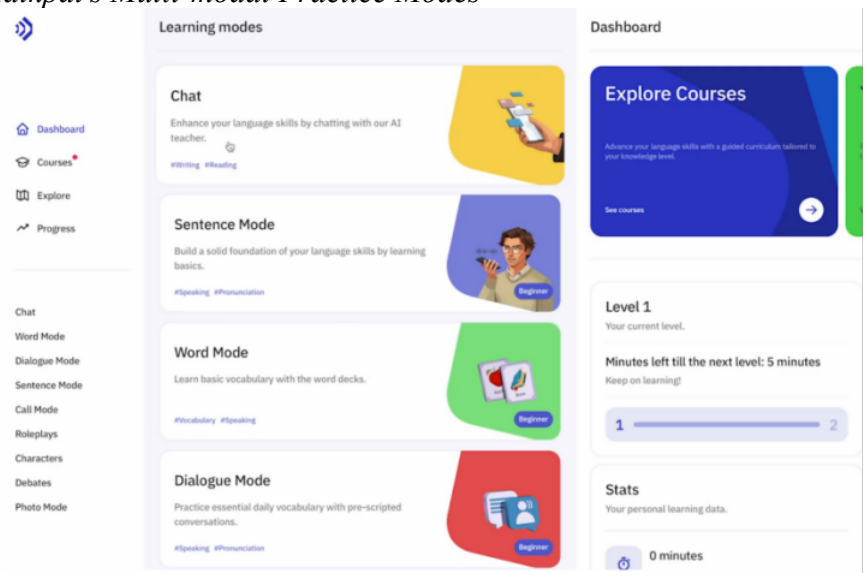
**Feature 1: GPT-Powered ConvAI for Multi-language Dialogue.** Talkpal uses a GPT-based conversational model to generate natural and context-sensitive dialogue across a wide range of languages and topics. Unlike scripted language learning applications, Talkpal allows open-ended interaction. Learners can discuss everyday topics, practice travel-related situations, or involve in more complex discussions depending on their proficiency level and learning goals (Padmanabhan, 2025). The platform's artificial intelligence can also adapt to the conversation content based on the learner's input, thereby creating a more personalized and responsive learning experience. The platform supports over 57 languages, which stands in stark contrast to the pedagogical approach of monolingual platforms such as Loora. Such extensive language coverage enables learners to undertake multilingual learning within a single system. Figure 2 shows examples of the languages available on Talkpal.

**Figure 2**
*Talkpal's Language Choices*




**Feature 2: Multi-modal Interactive Modes for Scenario Practice.** Talkpal has 9 innovative interaction modes, which are designed to cover different aspects of language learning. The Chat mode supports conversations based on general topics and provides grammar feedback. The Role-play mode simulates real-life situations such as ordering food, asking for directions, checking into a hotel, or attending a job interview. The Character Dialogue mode allows learners to interact with fictional or historical figures, making the exercises more vivid, involving and imaginative. Debate Mode encourages learners to express their views, construct arguments and employ more complex sentence structures. Call Mode focuses on speaking conversation practice, this helps to improve fluency and confidence in speaking. Furthermore, Sentence Mode, Dialogue Mode, Word Mode and Picture Mode provide more structured opportunities for grammar practice, vocabulary expansion and the use of descriptive language (Test Prep Insight, 2026). This multimodal framework provides learners with opportunities to repeatedly use language in context, which is crucial for developing communicative competence (Wang et al., 2017). It also maintains learner engagement by offering diverse content, reducing the monotony associated with repetitive digital exercises, and transforming practice into a more interactive experience. Figure 3 illustrates Talkpal's multimodal learning design, showcasing the various interaction modes available within a single platform.

**Figure 3**
*Talkpal's Multi-modal Practice Modes*



**Feature 3: Progress Tracking and Gamified Engagement.** Talkpal includes a progress tracking system, which makes the learning process more intuitive and quantifiable. The platform monitors metrics such as total study time, pronunciation performance, the number of exercises completed, the current number of consecutive days of study, and stage progress. It also provides feedback on grammar, vocabulary and pronunciation, helping learners identify specific strengths and weaknesses. This data-driven feedback helps to improve learners' self-awareness, as it encourages them to reflect on their learning patterns, track their progress and set more targeted goals. Gamified elements such as consecutive days of learning, badges, completed exercises and performance summaries also help to reinforce consistent learning habits (Julin, 2025). By visualizing learning progress, the platform encourages learners to monitor their own learning behavior and maintain a regular practice routine over the long term.

### *Limitations and Challenges*

Despite its strengths, Talkpal has several limitations. Firstly, it lacks a structured learning pathway. While the platform offers a variety of interactive modes, it does not always guide learners through a systematic course that includes progressive grammar, vocabulary and skills development (Mammadzada, C., 2026). This can pose a challenge for complete beginners who require clear explanations and step-by-step guidance. As a result, Talkpal functions more like a practice tool than a complete language learning system. Secondly, Talkpal's freemium model creates practical limitations. The free version's daily usage limit of approximately 10 minutes is insufficient for meaningful language learning. Although paid subscriptions unlock a wider range of features, price-sensitive users may compare Talkpal with free tools such as ChatGPT and question whether its additional gamified interface justifies the extra cost. Finally, Talkpal's extensive coverage of over 57 languages presents challenges in terms of quality control, as mainstream languages may receive stronger support than less commonly taught ones (Dunham, 2026). Consequently, Talkpal is best understood as a flexible supplementary tool within a blended learning mode, rather than a complete substitute for teachers or real-life communication.

### *Comparison Between Loora and Talkpal*

**Table 1**
*Comparison Between Loora and Talkpal*

| Dimension | Loora | Talkpal | Analysis |
|---|---|---|---|
| **Core Focus** | English speaking proficiency | Multi-language learning, 57+ languages | Loora emphasises depth in English speaking, whereas Talkpal prioritises breadth through multilingual coverage |
| **Technology Base** | Self-developed LLM and third-party LLMs | GPT powered platform | Both platforms rely on advanced LLM-based ConvAI to support interactive language learning |
| **Practice Mode** | Scenario-based dialogues, real-time feedback | Conversations, debates, character interactions, role-play | Talkpal offers greater mode diversity, while Loora focuses on structured English-speaking scenarios |

| | | | |
|---|---|---|---|
| **Customization Mode** | Adaptive difficulty and scenario customization (SwitchTools, 2026) | Level-based adaptation (Talkpal, 2025) | Both platforms provide customization, but Loora is more intelligent. It focuses on adjustment based on the learner's demonstrated proficiency, whereas Talkpal adapts based on user-selected language level. |
| **Free Tier** | Free trial for a certain period (e.g., 7 days) | Free plan, approximately 10 minutes per day | Talkpal is more accessible for initial users because it provides an ongoing limited free plan |
| **Target Users** | English learners, and professionals | Polyglots, diverse language learners, and casual users | The two platforms target different learner segments in the AI language learning market |
| **Strength** | Deep English-speaking practice and professional scenarios | Language variety and engaging practice modes | Loora's strength lies in specialized English practice, while Talkpal's strength lies in multilingual variety and interaction diversity |
| **Weakness** | Higher cost, English-only focus, limited mode variety | Potential depth limitations and audio quality issues | Loora charges more, only supports English and offers few practice modes. Talkpal lacks in-depth tutoring and has poor audio quality across its multilingual functions. (Loora, 2023; Talkpal, 2025) |

Table 1 provides a comprehensive comparison of Loora and Talkpal across eight dimensions. The comparison reveals that these two platforms occupy distinct positions within the AI-assisted language learning market. Loora is more specialized, focusing primarily on developing spoken English skills and professional communication practice; Talkpal, on the other hand, adopts a broader multilingual strategy, supporting over 57 languages and offering a wider variety of interactive practice modes. There are also some similarities between the two platforms. Table 2 summarizes the common strengths and limitations of these two tools.

**Table 2**
*Common Advantages and Limitations of Loora and Talkpal*

| Advantages | Limitations |
|---|---|
| • **Real-time feedback**<br>During the dialogue, the system provides vivid, immediate, and granular feedback to language learners by pointing out grammatical, pronunciation, or vocabulary errors in time to help them correct them immediately. | • **Lack of real interpersonal interaction**<br>The AI environment cannot replace the emotional support of human teachers, cultural resonance and non-verbal communication. |

| | |
|---|---|
| • **Multimodal interaction**<br>Support a variety of input and practice methods such as voice, text, dialogue mode to enrich users' learning experience.<br>• **Offer highly personalized and adaptive learning experiences**<br>Adjust the content of the dialogue according to the level goals of learners and the dynamics of historical interaction.<br>• **Break time and space restrictions**<br>Learners can conduct language training anytime and anywhere through mobile phones or computers without being bound by class time and place.<br>• **Reduce language learning costs**<br>Conv AI tools provide long-term exercises at a lower price than human tutors.<br>• **Progress tracking**<br>Automatically record the data of the accuracy rate of the course completion of the learning time and show the growth track in the form of visualization.<br>• **Exercise with low-pressure**<br>Practice repeatedly in an AI environment without judgment and social anxiety to reduce the psychological burden of speaking. | • **The restoration of cultural context is limited**<br>It is difficult for Loora and Talkpal to completely copy the differences between rational and humorous subtexts and regional culture in real communications.<br>• **Payment barrier**<br>Access to this high-quality feature often requires extra payment, posing a barrier for price-sensitive users.<br>• **Data privacy risk**<br>These two platforms will collect a large number of sensitive information in the habitual location of voice recordings. There is a hidden danger of leakage or abuse.<br>• **Insufficient continuity of dialogue**<br>In the on-site dialogue or cross-painting interaction, apps may forget the content discussed before, which leads to incoherent factors in the theme.<br>• **Limited depth**<br>For advanced learners or scenarios that requires in-depth writing training, the feedback provided is superficial. |

## Comparison Between ConvAI Tools and Human Teachers

Although tools such as Loora and Talkpal have demonstrated quantifiable benefits in terms of spoken practice, learner involvement and alleviating anxiety about speaking a foreign language, their effectiveness compared to human teachers remains to be verified. A systematic review conducted by Ji et al. (2022) examined 24 empirical studies on the application of ConvAI in language education. Their meta-analysis found that ConvAI primarily plays the roles of conversation partner, feedback provider, resource provider and needs analyzer. In contrast, human teachers function as designers, organizers, facilitators and emotional supporters. Consequently, the future development of ConvAI should focus on its role as a supplementary tool to human teachers. We believe such a collaboration is able to enhance teaching efficiency and encourage positive student involvement. Table 3 compares the effectiveness of ConvAI and human teachers across various pedagogical dimensions and proposes research ideas for future human-ConvAI collaborative teaching.

**Table 3**
*Comparative Analysis of ConvAI and Human Teachers*

| Dimension | Explanation | ConvAI | Human teacher | Future prospects of Human-ConvAI collaborations |
|---|---|---|---|---|
| **Collaborative problem solving** | Collaborative problem solving: encourages children to communicate effectively and fosters creativity and critical thinking (Aslan et al., 2025). | Monitor in real time, identify difficulties with collaboration and provide standardized intervention recommendations. | Focusing on high-value, flexible teaching decisions and emotional support | ConvAI can serve as a complementary tool to human teachers to help students build up collaborative problem-solving skills. It can provide real-time insights into group dynamics, and offer precise recommendations for interventions (Ji et al., 2022). |
| **Emotional connection** | Emotional Contagion Theory: People unconsciously mimic others' emotional expressions (Hatfield et al., 1993). | Using animated avatars, synthesized speech and virtual facial expressions to simulate joy. Provide a consistent positive emotional signal regardless of fatigue or mood. Furthermore, it can be programmed to avoid displaying negative expressions such as anger. (Peng et al., 2026) | Express genuine, natural emotions, communicate sincere care and empathy, and adapt their facial expressions in response to the real-time dynamics of the classroom. Teachers' emotional expressions have a direct impact on students' emotional experiences, and positive emotions help to foster a supportive learning environment (Becker et al., 2014). | ConvAI can act as an "emotional mirror" to detect students' emotional states and suggest or display appropriate positive expressions, thereby mitigating the negative effects of teacher burnout. Human teachers, meanwhile, can focus on conveying genuine emotional warmth. |

| | | | | |
|---|---|---|---|---|
| **Attention & cognitive engagement** | Joyful teachers are more successful in gaining and maintaining students' attention; positive emotions broaden cognitive focus and reduce anxiety (Marici et al., 2025). | It can be programmed to consistently display warm, professional and helpful traits, such as a patient tone and encouraging words (Zia, 2024). However, it lacks depth and consistency in long-term interactions (Yifan et al., 2025) | Through dynamic and engaging behavior such as moving around the classroom, varying their tone of voice and using humor. Through these, teachers can capture students' attention. They are also able to pick up on social cues that AI might overlook. | ConvAI can provide human teachers with real-time attention analysis alert, for example: "student X has been distracted for 20 seconds". Human teachers can then intervene in a targeted manner. Furthermore, in the future, AI may be able to autonomously implement "attention-grabbing" micro-interventions, whilst human teachers can focus on maintaining the teaching rhythm for the whole class (Zhang et al., 2025). |
| **Storytelling** | The narrator guides learners into contextualized tasks, facilitating task breakdown and providing immediate feedback (Khosrawi-Rad et al., 2025). | Function as the narrator to guide contextualized tasks and provide immediate prompts and feedback. | Design authentic, interdisciplinary storylines and guide students through role-play and emotional engagement (Kenshinbay et al., 2025). | ConvAI handles execution and feedback, whilst teachers are responsible for design and integration, creating a teaching model based on a pattern of "AI-driven narratives + teacher-led extension" (Ji et al., 2022) |

| | | | | |
|---|---|---|---|---|
| **Workload & efficiency** | Due to the ever-increasing workload, human teachers suffered from lower level of working efficiency (Aditya et al., 2026). | Answer frequently asked questions, help review learning materials, and record student performance. | Teachers are specialized in handling instructional design, classroom interaction, individual student support and professional development. | ConvAI acts as a "teaching assistant" to teachers, creating a division of labor – "AI handles structured tasks whilst teachers focus on creative teaching". At the same time, teachers can monitor and optimize the AI's behavior via a control panel (Aditya et al., 2026). |

In sum, ConvAI primarily plays the following roles in learning: dialogue partner and feedback provider, whilst human teachers are responsible for curriculum design, organizational arrangements and emotional support. Existing studies rarely provide direct comparisons between ConvAI and human teachers under the same instructional conditions. This gives us inspiration to collect data and understand the empirical differences between these two in future research.

## Conclusions and Implications

Through the comparative analysis of the two business platforms, Loora and Talkpal, this study evaluates the effectiveness and limitations of ConvAI in language learning and compares the difference between ConvAI tools and human teachers. The results of the study confirm that ConvAI tools provide measurable teaching benefits, which include scalable, immediate and multidimensional feedback, a low-anxiety environment conducive to oral practice, and reduced time and financial burdens of language exposure. These advantages are consistent with the propositions in the four proposed theories that ConvAI is an effective tool for language learning. Meanwhile both platforms carry drawbacks. Loora focuses on oral English teaching only, and some feedback are superficial for advanced learners. Pricing is also a concern for Loora. Talkpal provides a wide range of multilingual support with a lower price but suffered from Potential depth limitations and audio quality issues. More importantly, neither of these platforms can truly replicate the empathy, cultural intermediary or adaptive emotional support provided by human teachers. Our comparison between ConvAI tools and human teachers further reveals five major differences between them. This includes collaborative problem solving, emotional connection, attention and cognitive engagement, storytelling, and workload and efficiency. Our conclusion is that ConvAI may be a good tool in managing repetitive and feedback-intensive tasks, but human teachers still play an irreplaceable role in designing tasks, provide guidance and emotional support to language learners. ConvAI will not substitute human teachers, but researchers and practitioners should consider the way to well-coordinate division of labor between ConvAI and human educators. A collaborative pattern of ConvAI+human is the ideal learning mode for language learners in the future.

The present research carries several targeted implications for pedagogical research, educational practice, commercial application developers, and policy makers. First, in terms of pedagogical research, future work should focus more on the domain of ConvAI-human collaboration instead of ConvAI alone or human educator alone. Moreover, more empirical studies examining the learning effectiveness and learning outcomes should be conducted in this context. More detailed dimensions such as measurement of improvements in skill levels, learning motivation, cultural understanding and knowledge retention should be investigated. Secondly, ConvAI tools for educational practice should be embedded within a blended learning framework as supplementary resources. For example, Loora can be used to provide structured English-speaking practice with precise feedback, whilst Talkpal can be utilised for flexible multilingual exposure. It is recommended that teachers focus on higher-order activities, such as designing communicative tasks, cultural interpretation and emotional scaffolding, whilst utilising AI-generated analytical data to guide differentiated teaching. Third, developers must optimize ConvAI algorithms, making the app more natural and professional, and provide better support for educator's need. They should further prioritize improving the reliability of feedback, improving contextual memory, and implement robust data protection measures to build user trust and drive institutional adoption. Last, policymakers, especially those in the language schools, should ensure that relevant guidelines are given to both teachers and students in terms of the appropriate use of ConvAI in language learning. In addition, greater investment must be made in training both teachers and students in AI literacy to enhance their abilities to critically evaluate and effectively utilize these tools. At the institutional level, when selecting the right ConvAI tools for university-wide adoption, priority should be given to digital platforms that can demonstrate educational effectiveness, especially when collaborating with human teachers. By doing this, institutions can also drive the market's transition from single-function AI tutoring system to ConvAI-human collaborative teaching ecosystem.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that one generative AI tool – ChatGPT was used in drafting the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The authors further declare that, apart from this tool, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript.

## References

Aditya, D., Nagada, U., & Otermans, P. C. J. (2026). Redefining education: Addressing challenges with conversational AI. *Cogent Education*, *13*(1). https://doi.org/10.1080/2331186x.2025.2611649

Alieva, N. (2025). Human-AI communication: Sign interpretation challenges and contextual understanding in dialogue. *Journal of Universal Science Research*, *3*(4), 165–167.

Alkishri, W., Yousif, J. H., AI Husaini, Y., & Mahmood, A. (2025). Conversational AI in Education: A general review of Chatbot technologies and challenges. *Journal of Logistics Informatics and Service Science*, *12*(3), 264–282.

Aslan, S., Alyuz, N., Li, B., Durham, L. M., Shi, M., Sharma, S., & Nachman, L. (2025). An early investigation of collaborative problem solving in conversational AI-mediated learning environments. *Computers and Education: Artificial Intelligence*, *8*, 100393.

Becker, E. S., Goetz, T., Morger, V., & Ranellucci, J. (2014). The importance of teachers' emotions and instructional behavior for their students' emotions – An experience sampling analysis. *Teaching and Teacher Education*, *43*(43), 15–26. https://doi.org/10.1016/j.tate.2014.05.002

Chigbu, B. I., & Makapela, S. L. (2025). AI in education, sustainability, and the future of work: An integrative review of industry 5.0, education 5.0, and work 5.0. *Journal of Open Innovation: Technology*, *Market*, *and Complexity*, 100645.

Chua, H. W., & Annamalai, N. (2025). A bibliometric analysis and systematic review in AI chatbots in language teaching and learning. *International Journal of Technology in Education*, *8*(2), 274–331.

Dai, L., & Wu, F. (2025). An AI-powered conversational system for college students learning English as a second language. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13640-3

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Berlin: Springer Science & Business Media. https://doi.org/10.1007/978-1-4899-2271-7

Du, X. (2009). The affective filter in second language teaching. *Asian Social Science*, *5*(8), 162–165.

Dunham, N. (2026). *TalkPal Review: Is It Worth Subscribing in 2026?* https://languatalk.com/blog/talkpal-review/

Ellis, R. (2022). Task-based language teaching. In *Research Questions in Language Education and Applied Linguistics: A Reference Guide* (pp. 133–136). Springer Nature.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional Contagion. *Current Directions in Psychological Science*, *2*(3), 96–100. https://doi.org/10.1111/1467-8721.ep10770953

Ji, H., Han, I., & Ko, Y. (2022). A systematic review of conversational AI in language education: focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, *55*(1), 1–16. https://doi.org/10.1080/15391523.2022.2142873

Julin, J. (2025). *Gamification in Learning: Turning Work into Play*. https://craft-it4sd.eu/gamification-in-learning-turning-work-into-play/

Kenshinbay, T., Ghorbandordinejad, F., Bolatova, A., & Kiyassova, K. (2025). Enhancing language teacher reflection through narrative techniques: character development and critical incident analysis. *Bulletin of Language and Literature Studies*, *2*(1). https://doi.org/10.59652/blls.v2i1.482

Khosrawi-Rad, B., Keller, P. F., Benner, D., Grogorick, L., Borchers, A., Janson, A., Leimeister, J. M., & Robra-Bissantz, S. (2025). Promoting students' motivation in language education with gamified pedagogical conversational agents. *Computers & Education*, *238*, 105374. https://doi.org/10.1016/j.compedu.2025.105374

Kinginger, C. (2001). Sociocultural approaches to teaching and teachers' research. In R. Z. Lavine (Ed.), *Beyond the boundaries: Changing contexts in language learning* (pp. 201–225). McGraw-Hill Higher Education.

Krashen, S. D. (1986). *Principles and practice in second language acquisition*. Pergamon Press.

Loora. (2023). *Speak English with Loora AI*. https://www.loora.com/

Lunden, I. (2023). *Loora, a generative AI app that uses an audio interface to help users learn English, raises $9.25M.* https://techcrunch.com/2023/06/27/loora-a-generative-ai-app-that-uses-an-audio-interface-to-help-users-learn-english-raises-9-25m/

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, *82*(4), 545–562.

Mammadzada, C. (2026). *Best AI Language Learning App 2026: Top AI Language Learning Apps Ranked*. https://www.enverson.com/what-are-the-best-ai-language-learning-apps-2026-edition

Maricheva A., & Kyrychenko, R. (2026). *What is ASR: Understanding Automatic Speech Recognition*. https://softteco.com/blog/what-is-asr

Marici, M., Iosim, I., & Marin, C. D. (2025). The role of teachers' emotional facial expressions on student perceptions and engagement for primary school students-an experimental investigation. *Frontiers in Psychology*, *16*, 1613073.

Padmanabhan, I. (2025). *Honest Talkpal Review: Is Talkpal AI Worth it (When ChatGPT is Free)?* https://ilampadmanabhan.medium.com/honest-talkpal-review-is-talkpal-ai-worth-it-when-chatgpt-is-free-sep-2025-b805fabacfb8

Peng, X., Qian, Y., & He, Y. (2026). When AI chatbots understand emotions: exploring the mechanisms of self-disclosure from emotional arousal to psychological acceptance-A hybrid SEM-ANN-NCA analysis across multiple interaction configurations. *Frontiers in Psychology*, *16.* https://doi.org/10.3389/fpsyg.2025.1724313

PracticeMe. (2026). *Loora AI review: Honest take on English practice*. https://practiceme.app/vs/loora

Shadiev, R., Kaliyeva, A., Kairatova, A., Yerbulatova, Z., Ryskulbek, A., & Beisembayeva, Z. (2025). Exploring the role of conversational AI in developing communicative skills: A systematic review. *Digital Applied Linguistics*, *4*, 103058.

Stryker C., & Holdsworth, J. (2026). *What is NLP (natural language process)?* https://www.ibm.com/think/topics/natural-language-processing

SwitchTools. (2026). *Loora*. https://www.switchtools.io/tool/loora

Talkpal. (2025). *Talkpal AI Language Learning: Official Website and Pricing Information*. https://talkpal.ai/

Test Prep Insight. (2026). *Talkpal Review Our In-depth Review of TalkPal*. https://testprepinsight.com/reviews/talkpal-review/

Wang, Y. F., Petrina, S., & Feng, F. (2017). VILLAGE—virtual immersive language learning and gaming environment: immersion and presence. *British Journal of Educational Technology*, *48*(2), 431–450.

Weizenbaum, J. (1966). ELIZA-A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45.

Wu, L., Zheng, Z., Qiu, Z., Wang, H., Gu, H., Shen, T., Qin, C., Zhu, C., Liu, Q., Xiong, H., & Chen, E. (2024). A survey on large language models for recommendation. *World Wide Web*, *27*(5), 60.

Xing, C., & Saeed, M. A. (2025). A Systematic review on artificial intelligence (AI) technologies in ESL/EFL speaking skills. *International Journal of TESOL Studies*, *8*(2), 240–270.

Yifan, L., Kangping, X., Yanzhen, L., Yang, Y., & Andrew, Y. (2025). *NeurIPS Probing the Lack of Stable Internal Beliefs in LLMs.* https://neurips.cc/virtual/2025/loc/mexico-city/129953

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, *16*(1), 1–27.

Zhang, H., Su, Y., Xu, X., Lei, V. N.-L., & Hao, S. (2025). Beyond theory: A mixed-methods exploration of postgraduate engagement with ChatGPT for English speaking proficiency. *International Journal of Technology in Education*, *8*(4), 1022–1047. https://doi.org/10.46328/ijte.1221

Zhang, J., Beckmann, N., & F. Beckmann, J. (2018). To talk or not to talk: A review of situational antecedents of willingness to communicate in the second language classroom. *System*, *72*, 226–239.

Zia, K. (2024). Chatbots as conversational partners: their effectiveness in facilitating language acquisition and reducing foreign language anxiety. *Journal of Applied Linguistics and TESOL (JALT)*, *7*(4), 238–255. https://jalt.com.pk/index.php/jalt/article/view/38

**Contact email:** cong.qi@polyu.edu.hk

# Children as Guides in Storytelling: Exploring Imagination and Expression in Child–AI Collaborative Narrative Play

Ting Xu, The Hong Kong University of Science and Technology, China
Xian Xu, Lingnan University, Hong Kong

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

The rapid adoption of generative artificial intelligence (AI) in education does not automatically enhance children's creativity. Prevailing prompt–response interaction patterns often position children as passive consumers, risking the homogenization of their distinctive imagination and narrative style. This study proposes and examines child-guided co-narration, an alternative paradigm that repositions the child as narrative authority and the AI as a responsive partner to children's unstructured, play-based expressions, such as spontaneous drawing, block building, and improvised telling. We instantiated the paradigm in StorySketcher, a multi-agent AI system that transforms children's sketches and voices into animated, multilingual stories while preserving creative agency through Socratic prompting. A two-phase mixed-methods study combined participatory design (n = 12 children, aged 5–8) with a four-week quasi-experimental educational study (n = 36 children, aged 6–8) comparing the child-guided system with a prompt–response AI storytelling application. Narrative agency was operationalized as the child-initiated decision ratio; narrative complexity through story-grammar and cohesion coding; originality through consensual expert assessment; and dynamic visual literacy through a purpose-built rubric. Results show that the child-guided condition produced significantly higher narrative agency, more complex and more original stories, and stronger dynamic visual literacy. We argue that treating the child as guide reframes AI from ghostwriter to an extension and mirror of children's imagination, offering a framework for protecting and enhancing children's artistic and narrative expression in the AI era.

*Keywords:* child–AI interaction, generative AI in education, narrative development, co-creative storytelling, visual literacy, participatory design

iafor
The International Academic Forum
www.iafor.org

## Introduction

Storytelling is one of the earliest and most powerful ways in which children organize experience, rehearse social understanding, and externalize imagination (Bruner, 1986; Paley, 1990). Long before they can write fluently, young children tell stories through drawings, gestures, objects, and improvised speech, and these multimodal narratives constitute a foundational form of cultural participation and meaning-making (Vygotsky, 1978, 2004). How emerging technologies mediate this activity is therefore not a peripheral design question but a question about the conditions under which children's imagination develops.

Generative artificial intelligence (AI) has rapidly entered early learning environments in the form of story generators, drawing-to-image tools, and conversational agents (Su & Yang, 2022; UNESCO, 2023). The dominant interaction pattern of these tools is prompt–response: the user issues an instruction, and the system returns a polished artifact. When this pattern is transplanted into early childhood, it tends to reposition the child from storyteller to requester, and ultimately to consumer, of machine-generated narrative. Recent evidence suggests that although generative AI can raise the mean quality of individual creative products, it simultaneously narrows the collective diversity of ideas (Doshi & Hauser, 2024). For young children, whose narrative voices are still forming, such homogenization is not merely an aesthetic loss; it threatens the developmental process through which a distinctive imagination is practiced and consolidated (Vygotsky, 2004).

We argue that the central design question is not whether children can produce prettier stories with AI, but who is telling whose story. This paper proposes child-guided co-narration, a human–AI paradigm in which the child remains the narrative authority—the one who invents characters, decides events, and resolves ambiguity—while the AI is architecturally constrained to respond to the child's unstructured, play-based expressions: spontaneous sketches, block constructions, and improvised telling. We instantiate the paradigm in StorySketcher, a multi-agent system that transforms children's drawings and voices into animated, multilingual stories, using Socratic prompts that return narrative decisions to the child rather than supplying content.

We report a two-phase mixed-methods study that combines participatory design with an educational experiment, addressing three research questions: (RQ1) How does the child-guided paradigm shape children's narrative agency during child–AI co-narrative play? (RQ2) How does it influence the complexity and originality of children's narrative expression? (RQ3) How does playful, child-guided interaction support children's dynamic visual literacy? The contributions of this work are threefold: (1) a conceptual framework that reframes child–AI storytelling around narrative initiative rather than output quality; (2) StorySketcher, a working multi-agent instantiation of the framework; and (3) empirical evidence from 48 children showing measurable gains in narrative agency, narrative complexity, originality, and dynamic visual literacy.

## Related Work

### Narrative, Play, and Drawing in Early Childhood

Developmental theory has long treated play as the engine of imagination. For Vygotsky (1978, 2004), children's play is not the recycling of experience but its creative recombination: imagination arises precisely when the child detaches meaning from the immediate object and acts within an imagined situation. Piaget (1962) similarly described symbolic play as the child's

assimilation of the world to the self, and Bruner (1986) identified narrative as one of the two fundamental modes through which human beings construct reality. Classroom research has shown the pedagogical power of this view: Paley's (1990) storytelling-and-story-acting practice demonstrated that when children dictate and enact their own stories, narrative becomes a vehicle for community, literacy, and moral reasoning. Drawing occupies a special place in this ecology. Levin and Bus (2003) showed that young children's emergent writing is literally based on drawing, and analyses of children's spoken narratives (Peterson & McCabe, 1983) provide mature coding schemes for the structural complexity of what children tell. Taken together, this literature implies that any technology for children's storytelling should protect the child's initiative and meet the child in play-based, multimodal expression rather than in formal text.

## Digital and AI-Supported Storytelling for Children

Interaction design for children's storytelling predates generative AI. StoryMat created a play space that recorded and replayed children's collaborative fantasy play (Ryokai & Cassell, 1999); Fiabot! supported school-based digital storytelling on mobile devices (Rubegni & Landoni, 2014); and constructionist environments such as ScratchJr and programmable toys framed coding itself as a playful, expressive medium (Bers, 2018; Resnick, 2017). With large language and image models, a new generation of systems has emerged. StoryBuddy supports parent–child interactive storytelling with adjustable parental involvement (Zhang, Xu, et al., 2022); StoryDrawer pairs a conversational agent with children's drawings to scaffold creative visual storytelling (Zhang, Yao, et al., 2022); and Mathemyths uses child–AI co-creative storytelling to teach mathematical language (Zhang et al., 2024). These systems demonstrate clear learning and engagement benefits, and they increasingly position AI as a partner rather than a tool.

## The Agency Gap: From AI-Led Generation to Child-Guided Co-narration

Across this literature, however, three gaps remain. First, initiative asymmetry persists: even in co-creative systems, the AI typically proposes content—plot continuations, character actions, image completions—that the child selects or edits, so the locus of invention still sits substantially with the machine. Second, evaluation practice mirrors this asymmetry: systems are chiefly evaluated on the quality, engagement, or learning value of their outputs, rarely on the distribution of narrative initiative during the process. Third, key outcomes invoked in debates about children's creativity with AI—agency, originality, and visual literacy—are seldom operationalized with validated measures, making claims difficult to compare or verify. These gaps matter because evidence from adult creativity research indicates that generative AI raises average output quality while reducing collective novelty (Doshi & Hauser, 2024), and childhood is precisely the period in which an original narrative voice is formed (Vygotsky, 2004). Child-computer interaction scholarship has long insisted that children should be design partners and protagonists, not end users (Druin, 2002; Hourcade, 2015), and multiliteracies theory reminds us that meaning is made across modes—image, movement, and speech as much as writing (Kress & van Leeuwen, 2006; The New London Group, 1996). Visual literacy itself, defined as the ability to interpret and compose meaning through visual signs (Avgerinou & Ericson, 1997), acquires a dynamic dimension when children author for animated media. The child-guided paradigm responds directly to these gaps: it asks what changes—for agency, expression, and visual literacy—when AI is designed to follow rather than to lead.

# The StorySketcher System

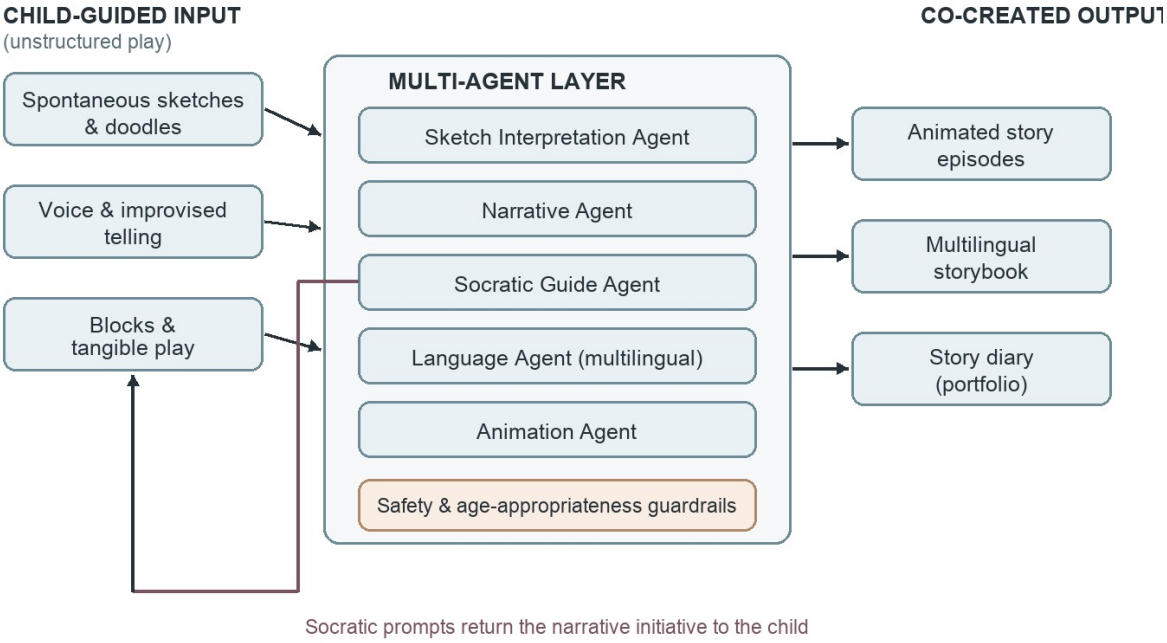
## Design Principles

StorySketcher was designed around three principles. (1) The child leads; the AI responds. The system contains no text box in which a child is asked to prompt for a story; narrative content is always elicited from the child's own drawing, speech, or play. (2) Play is the input modality. Children's spontaneous sketches, block arrangements, and improvised telling are treated as legitimate authorial acts, echoing constructionist and multiliteracies perspectives (Bers, 2018; The New London Group, 1996). (3) Output must be safe, revisable, and multilingual. Every generated element is filtered for age-appropriateness, remains editable by the child, and can be retold in the child's home language or a target foreign language, supporting both identity expression and language exposure.

## Multi-agent Architecture

StorySketcher is implemented as a team of cooperating agents rather than a single generative model (Figure 1). The Sketch Interpretation Agent converts a child's drawing into a candidate scene description, which is always spoken back to the child for confirmation or correction (for example, “I see a house by the sea and a small cat—did I get it right?”). The Narrative Agent maintains the emerging story world: characters, settings, and events endorsed by the child. The Socratic Guide Agent monitors the story for open possibilities and poses non-directive questions (“What does the cat do when the rain starts?”) that return the next decision to the child; it never supplies plot content. The Language Agent renders the endorsed story in selectable languages at an age-appropriate register, and the Animation Agent turns confirmed scenes into short animated episodes that preserve the child's original strokes, colors, and composition. A safety and age-appropriateness layer governs all generated text and imagery. Crucially, the architecture enforces the paradigm technically: agents may interpret, ask, translate, and animate, but only the child can introduce narrative facts.

**Figure 1**
*The StorySketcher Multi-agent Architecture. Agents Interpret, Question, Translate, and Animate; Narrative Initiative Remains With the Child.*

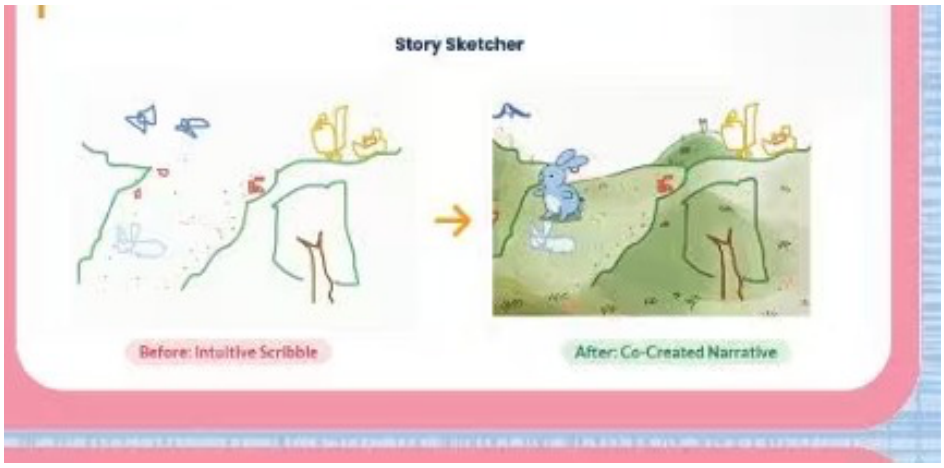



## Interaction Flow

A typical episode proceeds as a loop. The child draws or builds freely (Figure 2); the system reflects its interpretation aloud; the child confirms, corrects, or extends it; the Socratic Guide offers one open question; and the child answers in speech or with new strokes. Confirmed scenes accumulate in a Story Diary, which the child can revisit, reorder, and narrate. On request, scenes are animated and the story is retold in a chosen language, so a single narrative can move between the child's first language and a target language while remaining recognizably the child's own. Figure 3 shows a characteristic transformation: the child's intuitive scribble is preserved in composition and line, while co-created narrative detail is added around it. Over multiple sessions, episodes connect into a continuous animation or a printed picture book.

**Figure 2**
*The StorySketcher Workspace: Drawing Canvas (Center), Story Diary (Upper Left), and the Confirmation Dialogue (Lower Left)*

**Figure 3**
*From Scribble to Story: A Child's Intuitive Drawing (Left) and the Co-created Narrative Scene (Right), Preserving the Child's Composition*




## Methodology

### Research Design

We adopted a two-phase mixed-methods design (Creswell & Plano Clark, 2018). Phase 1 used participatory design to build and refine StorySketcher with children and to develop the coding instruments used later. Phase 2 was a quasi-experimental educational study comparing the child-guided system with a prompt–response AI storytelling application. Quantitative and qualitative strands were collected concurrently in Phase 2 and integrated through a joint display, so that behavioral logs, rated artifacts, and children's own accounts could corroborate one another.

### Phase 1: Participatory Design

Twelve children (six girls, six boys; aged 5–8, M = 6.6, SD = 0.9) from one kindergarten and one primary school in Guangzhou, China, participated in two iterative design cycles over six weeks (Table 1). Following cooperative inquiry practice (Druin, 2002), children worked as design partners in 45-minute sessions: they drew, told stories, reacted to paper and digital prototypes, and articulated what the machine should and should not be allowed to do. Session videos, field notes, and design artifacts were analyzed thematically (Braun & Clarke, 2006). Three outcomes directly shaped the final system: children strongly resisted machine-invented plot twists ("that's not my story"), confirming the need for architectural rather than merely stylistic constraint on the AI; children treated the system's verbal reflection of their drawings as a game, which motivated the confirmation loop; and children repeatedly switched languages for dramatic effect, motivating the Language Agent.

### Phase 2: Educational Experiment

Thirty-six children (18 girls; aged 6–8, M = 7.1, SD = 0.8) who had not participated in Phase 1 were recruited from two intact classes of a primary school in Guangzhou; classes were randomly assigned to condition to avoid treatment diffusion. Over four weeks (eight 40-minute sessions, two per week), the child-guided group (n = 18) created stories with StorySketcher, while the comparison group (n = 18) used a prompt–response AI storytelling application

representative of mainstream tools, in which the child enters or selects a prompt and the system generates a complete story with illustrations. Both groups addressed the same weekly story themes, with the same teachers and session structure. Before and after the intervention, each child completed an individual picture-prompt storytelling task ("tell me the story of these pictures") without any technology; stories were audio-recorded and transcribed verbatim.

**Table 1**
*Participant Characteristics Across the Two Study Phases*

| Phase / Group | n | Girls | Age range | Age M (SD) | Setting |
|---|---|---|---|---|---|
| Phase 1: participatory design | 12 | 6 | 5–8 | 6.6 (0.9) | Kindergarten + primary |
| Phase 2: child-guided | 18 | 9 | 6–8 | 7.2 (0.8) | Primary school |
| Phase 2: comparison | 18 | 9 | 6–8 | 7.0 (0.8) | Primary school |

## Measures

Four constructs were operationalized (Table 2). Narrative agency was measured as the child-initiated decision ratio: the proportion of narrative facts (characters, settings, events, and resolutions) in the final artifact that originated with the child rather than the AI, coded from session logs and screen recordings by two trained coders (Cohen's $\kappa = .86$). Narrative complexity was scored on the pre- and post-test transcripts using an eight-element story-grammar scheme adapted from Peterson and McCabe (1983), plus causal cohesion (causal connectives per ten clauses) and vocabulary diversity (moving-average type–token ratio). Originality of the final stories was rated with the Consensual Assessment Technique (Amabile, 1982): three early-childhood education experts, blind to condition, rated each story on a six-point scale (intraclass correlation, ICC(2,3) = .84). Dynamic visual literacy was assessed with a purpose-built rubric covering sequential representation, depiction of motion and change, symbolic annotation, and image–narrative correspondence (0–3 per dimension; total 0–12), validated in a pilot (Cronbach's $\alpha = .88$). Engagement and experience were captured through session observation, a five-point smileyometer after each session, and brief semi-structured interviews with children and teachers.

**Table 2**
*Operationalization of Key Constructs, Data Sources, and Analyses*

| Construct | Operationalization | Data source | Analysis |
|---|---|---|---|
| Narrative agency | Child-initiated decision ratio (child- vs. AI-originated narrative facts) | Session logs, recordings | t-test; thematic coding |
| Narrative complexity | Story-grammar score (0–8); causal cohesion; vocabulary diversity | Pre/post story transcripts | ANCOVA (pre as covariate) |
| Originality | Consensual Assessment Technique, 3 blinded experts, 1–6 scale | Final stories | t-test; ICC |
| Dynamic visual literacy | 4-dimension rubric (0–12) | Final animated stories | t-test; rubric scoring |

### Data Analysis

Quantitative outcomes were analyzed with independent-samples t-tests for process measures and ANCOVA (pre-test as covariate) for post-test story measures, with effect sizes reported as Cohen's d and partial $\eta^2$. Qualitative data from Phase 1 and Phase 2 interviews, observations, and open child comments were analyzed thematically (Braun & Clarke, 2006). Integration occurred in a joint display mapping statistical results onto explanatory themes.

### Ethics

The study was approved by the university's human research ethics committee. Parents provided written informed consent and children gave age-appropriate assent, with the right to withdraw at any time. All AI processing for the study ran on institutionally controlled infrastructure with data minimization; no child data were used for model training; and all generated content passed the system's safety layer and teacher review, consistent with emerging international guidance on generative AI in education (UNESCO, 2023).

## Results

### Narrative Agency (RQ1)

Children in the child-guided condition originated, on average, 71% of the narrative facts in their final stories ($M = .71$, $SD = .14$), compared with 42% in the prompt–response condition ($M = .42$, $SD = .15$), a large and significant difference, $t(34) = 5.99$, $p < .001$, $d = 2.00$ (Table 3). The process data reveal what this ratio means behaviorally. In the comparison condition, children typically accepted the first generated story and edited at the margins; in the child-guided condition, children routinely rejected or corrected the system's interpretations ("No! The cat is not scared, she is brave!"), treating the AI as a fallible assistant whose job was to keep up with their intentions.

## Narrative Complexity and Originality (RQ2)

Controlling for pre-test scores, the child-guided group produced post-test stories with significantly higher story-grammar complexity (M = 6.6, SD = 0.9) than the comparison group (M = 5.1, SD = 1.2), $F(1, 33) = 17.84$, $p < .001$, partial $\eta^2 = .35$, and greater causal cohesion (M = 3.8 vs. 2.6 connectives per ten clauses), $t(34) = 3.78$, $p = .001$, $d = 1.26$. Vocabulary diversity did not differ significantly, $t(34) = 1.99$, $p = .055$, suggesting that gains reflect structural ambition rather than surface elaboration. Blinded expert ratings of originality were significantly higher for child-guided stories (M = 4.6, SD = 0.8) than comparison stories (M = 3.7, SD = 0.9), $t(34) = 3.17$, $p = .003$, $d = 1.06$. Raters' comments indicated that comparison stories converged on similar plots and stock characters, whereas child-guided stories contained idiosyncratic premises traceable to each child's drawings—a pattern consistent with homogenization concerns in the adult creativity literature (Doshi & Hauser, 2024).

## Dynamic Visual Literacy (RQ3)

Child-guided stories scored significantly higher on the dynamic visual literacy rubric (M = 9.3, SD = 1.5) than comparison stories (M = 6.2, SD = 1.7), $t(34) = 5.80$, $p < .001$, $d = 1.93$. The largest differences appeared in depiction of motion and change and image–narrative correspondence: because children had to make their drawn worlds move, they iteratively revised sequences of sketches, added motion marks and symbols, and negotiated with the system how a static drawing could become an event—practices close to the grammar of visual design (Kress & van Leeuwen, 2006).

**Table 3**
*Quantitative Outcomes for the Child-Guided and Comparison Conditions*

| Outcome | Child-guided M (SD) | Comparison M (SD) | Test | Effect size |
|---|---|---|---|---|
| Child-initiated decision ratio | .71 (.14) | .42 (.15) | $t(34) = 5.99$, $p < .001$ | $d = 2.00$ |
| Story-grammar score (post) | 6.6 (0.9) | 5.1 (1.2) | $F(1, 33) = 17.84$, $p < .001$ | $\eta^2 = .35$ |
| Causal cohesion (per 10 clauses) | 3.8 (0.9) | 2.6 (1.0) | $t(34) = 3.78$, $p = .001$ | $d = 1.26$ |
| Vocabulary diversity | .72 (.04) | .69 (.05) | $t(34) = 1.99$, $p = .055$ | — |
| Originality (CAT, 1–6) | 4.6 (0.8) | 3.7 (0.9) | $t(34) = 3.17$, $p = .003$ | $d = 1.06$ |
| Dynamic visual literacy (0–12) | 9.3 (1.5) | 6.2 (1.7) | $t(34) = 5.80$, $p < .001$ | $d = 1.93$ |

## Children's Experience

Thematic analysis of interviews and observations produced three themes. First, ownership: children described the stories as unmistakably theirs ("The AI is my helper. I tell the story, it draws with me," P17, age 7), and teachers noted that children defended their plots against

machine suggestions. Second, drawing as thinking: children used new strokes to answer the Socratic Guide's questions, visibly reasoning through the pen ("First I didn't know what the fox does. Then I draw—ah, the fox opens an umbrella!" P09, age 6). Third, from teller to director: children issued cinematographic instructions ("make the wave bigger," "say it in English for grandma"), orchestrating animation and language as expressive resources. Engagement remained high across sessions (mean smileyometer rating 4.5 of 5), and no child withdrew.

## Discussion

### From Ghostwriter to Respondent

The results support the paper's central claim: when the locus of narrative initiative is moved from the model to the child—architecturally, not merely rhetorically—children's agency, structural ambition, and originality rise together. Existing child–AI storytelling systems have progressively enriched the partner role of AI (Zhang, Xu, et al., 2022; Zhang, Yao, et al., 2022; Zhang et al., 2024), but the AI in these systems still proposes content. StorySketcher demonstrates a complementary position: an AI whose generative capacity is deliberately bounded by the child's play. In Bandura's (2001) terms, the system preserves the child's authorship of intentions while the machine supplies executional power.

### Imagination Extended, Not Replaced

Vygotsky (2004) warned that imagination develops through the child's own combinatorial activity, not through passive reception of finished images. Our findings give this warning contemporary operational form. The comparison condition, which produced objectively polished stories, showed convergence toward stereotyped plots—the child-level analogue of the diversity loss documented by Doshi and Hauser (2024). The child-guided condition, by contrast, externalized and amplified idiosyncrasy: the AI acted as a mirror that made each child's imagination visible, discussable, and revisable, in line with constructionist accounts of learning through personally meaningful artifacts (Bers, 2018; Resnick, 2017). The significant gains in dynamic visual literacy further suggest that authorship for animated media cultivates a genuinely multimodal competence anticipated by multiliteracies theory (The New London Group, 1996).

### Implications for Design and Pedagogy

For designers, the study offers a transferable pattern: constrain generative systems so that they may interpret, question, translate, and render—but not invent—on behalf of young learners; make the system's interpretations inspectable and correctable; and log the distribution of initiative as a first-class evaluation metric. For educators, StorySketcher sessions functioned as a studio pedagogy in which drawing, telling, language switching, and directing were woven into one activity, requiring little prompting beyond the Socratic questions the system itself posed. Two extensions follow naturally. First, the Socratic Guide could evolve into a fully conversational partner that sustains extended dialogue about characters' motives and alternative endings, bringing conversational AI into early childhood settings under the same child-led constraint; this is a promising direction for deepening narrative reflection. Second, the paradigm should be tested in homes and informal learning settings, where parent mediation may further shape how children negotiate authorship with AI. We concur with our reviewers' intuition that the child-as-guide principle deserves broad advocacy: it offers a concrete,

evaluable answer to the question of what child-centered generative AI should look like in the AI era.

## Limitations and Future Work

This study has several limitations. The sample was modest and drawn from one urban region; intact-class assignment risks cluster effects; and the four-week intervention cannot reveal long-term developmental consequences or novelty effects. The comparison tool, while representative of mainstream prompt–response products, was not a commercial system optimized over years of use. Measures of agency, though reliable, were labor-intensive to code. Future work will pursue longitudinal and cross-cultural replications, the conversational extension of the Socratic Guide, home-based deployment, and automated initiative analytics that would allow the child-guided principle to be evaluated—and enforced—at scale.

## Conclusion

This paper introduced child-guided co-narration, a paradigm that casts the child as narrative authority and AI as a responsive partner to children's play-based expression, and instantiated it in StorySketcher, a multi-agent system that turns sketches and voices into animated, multilingual stories. A two-phase mixed-methods study showed that, relative to a prompt–response AI tool, the child-guided system substantially increased children's narrative agency, narrative complexity, originality, and dynamic visual literacy. The deeper contribution is conceptual: in an era when generative systems can tell flawless stories on demand, the educational task is to ensure that the stories children grow up with remain theirs. Designing AI to follow the child's lead is a practical, measurable way to protect and extend children's imagination and expression.

## Acknowledgements

We thank the participating children, families, and teachers for their creativity and trust, and our research assistants for coding support. An early prototype of StorySketcher received a Gold Medal at the 51st International Exhibition of Inventions Geneva.

## References

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology, 43*(5), 997–1013.

Avgerinou, M., & Ericson, J. (1997). A review of the concept of visual literacy. *British Journal of Educational Technology, 28*(4), 280–291.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1–26.

Bers, M. U. (2018). *Coding as a playground: Programming and computational thinking in the early childhood classroom*. Routledge.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Bruner, J. S. (1986). *Actual minds, possible worlds*. Harvard University Press.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances, 10*(28), eadn5290.

Druin, A. (2002). The role of children in the design of new technology. *Behaviour & Information Technology, 21*(1), 1–25.

Hourcade, J. P. (2015). *Child-computer interaction*. Author.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.

Levin, I., & Bus, A. G. (2003). How is emergent writing based on drawing? Analyses of children's products and their sorting by children and mothers. *Developmental Psychology, 39*(5), 891–905.

The New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review, 66*(1), 60–92.

Paley, V. G. (1990). *The boy who would be a helicopter: The uses of storytelling in the classroom*. Harvard University Press.

Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics: Three ways of looking at a child's narrative*. Plenum Press.

Piaget, J. (1962). *Play, dreams and imitation in childhood*. W. W. Norton.

Resnick, M. (2017). *Lifelong kindergarten: Cultivating creativity through projects, passion, peers, and play*. MIT Press.

Rubegni, E., & Landoni, M. (2014). Fiabot! Design and evaluation of a mobile storytelling application for schools. In *Proceedings of the 2014 Conference on Interaction Design and Children* (pp. 165–174). ACM.

Ryokai, K., & Cassell, J. (1999). StoryMat: A play space for collaborative storytelling. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 272–273). ACM.

Su, J., & Yang, W. (2022). Artificial intelligence in early childhood education: A scoping review. *Computers and Education: Artificial Intelligence, 3*, 100049.

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Vygotsky, L. S. (2004). Imagination and creativity in childhood. *Journal of Russian and East European Psychology, 42*(1), 7–97.

Zhang, C., Liu, X., Ziska, K., Jeon, S., Yu, C., & Xu, Y. (2024). Mathemyths: Leveraging large language models to teach mathematical language through child–AI co-creative storytelling. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. ACM.

Zhang, C., Yao, C., Wu, J., Lin, W., Liu, L., Yan, G., & Ying, F. (2022). StoryDrawer: A child–AI collaborative drawing system to support children's creative visual storytelling. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. ACM.

Zhang, Z., Xu, Y., Wang, Y., Yao, B., Ritchie, D., Wu, T., Yu, M., Wang, D., & Li, T. J.-J. (2022). StoryBuddy: A human-AI collaborative chatbot for parent-child interactive storytelling with flexible parental involvement. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1–21). ACM.

**Contact email:** txu560@connect.hkust-gz.edu.cn

# A Competency-Oriented System of Chemistry Tasks for Grade 7: Scientific and Methodological Foundations and Empirical Validation of Effectiveness

Indira Aimbetova, Akhmet Yassawi International Kazakh-Turkish University, Kazakhstan
Dinara Berdi, Akhmet Yassawi International Kazakh-Turkish University, Kazakhstan
Akylbek Meirbekov, Akhmet Yassawi International Kazakh-Turkish University, Kazakhstan

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

The article examines the scientific and methodological principles underlying the development and implementation of a task system aimed at fostering functional literacy within the Grade 7 chemistry curriculum. The relevance of the study is *обусловлена* limited laboratory infrastructure in many Kazakhstani schools, especially in rural areas, as well as insufficient integration of digital and interactive learning tools. These constraints reduce opportunities for experimental work and hinder the development of research skills and scientific literacy required by modern educational standards and international assessments such as PISA and TIMSS. The study aims to design and pedagogically validate a system of contextual, experimental, problem-based, and analytical tasks that connect theoretical chemical concepts with real-life situations. The methodological framework includes curriculum and document analysis, comparative pedagogical analysis, task classification, and elements of inquiry-based and competence-oriented learning approaches. The empirical component involved surveys of 120 school students and 120 undergraduate and master's students enrolled in chemistry-related programs. The findings indicate that systematic use of functionally oriented tasks increases students' interest in chemistry by 32–47%, enhances analytical and experimental skills, and supports the development of scientific thinking and environmental awareness. The results demonstrate that the proposed task system significantly improves the quality of chemistry instruction at the lower secondary level while serving as an effective tool for developing functional literacy and research competencies. The developed instructional package has strong practical applicability and can be adapted for use across diverse educational contexts.

*Keywords:* chemical education, Mozabook, methodical complex, critical evaluation, classroom tasks

iafor
The International Academic Forum
www.iafor.org

## Introduction

The extent to which digital technologies such as Web 2.0 tools, podcasts, blogs, social media, email, web forums, online chats, electronic dictionaries, online informaation and reference resources, and videoconferencing – are used in chemistry education in Kazakhstan remains comparatively lower than in developed countries. In school chemistry education, it is essential to preserve the educational triad of "knowledge-motivation and values-competencies and skills." Such an approach contributes to the development of students' foreign-language communicative competence, promotes independent learning, and expands their vocabulary through comments, discussions, and debates, thereby fostering an aspiration for lifelong learning.

A substantial body of reputable international research emphasizes that digital technologies enable multimedia content to be created, continuously updated, and systematically organized. Digital platforms can incorporate audio, video, and graphical materials, including authentic foreign-language texts, podcasts, and video excerpts. These resources are particularly valuable for students studying chemistry. Presenting chemical processes in a variety of formats allows learners to observe them directly and facilitates mastery of the chemistry curriculum by strengthening visual and photographic memory.

A considerable number of pedagogical studies in Kazakhstan have examined the benefits of electronic educational publications. Furthermore, several studies have developed methods for using these resources in chemistry education. Their findings indicate that electronic educational content has a highly positive influence on chemistry learning. A sufficient body of research has also demonstrated that the principal objective of educational modernization is to improve the quality of education in ways that promote the harmonious development of the individual and cultivate the skills required to live effectively in an increasingly information-rich society. For example, one group of studies examines AI-supported, gamified, and intelligent learning environments in school education. These studies provide valuable insights into how adaptive tools can personalize learning, sustain student engagement, and support the digital transformation of science classrooms (Grankina et al., 2026; Niyazova et al., 2026).

Information systems play a particularly important role in the digitalization of educational processes. Such systems should enable school administrators to manage educational institutions effectively, allow teachers and educators to monitor students' academic progress systematically, and provide parents with timely information about their children's achievement, attendance, and engagement in learning. Moreover, information systems strengthen communication among participants in the educational process – including school administrators, teachers, educators, and students – and parents, while facilitating the rapid exchange of essential information. In this way, they create conditions for parents to participate actively in school life, maintain regular communication with teachers, and collaborate in supporting students' education.

The primary objective of the ongoing reforms in Kazakhstan's education sector is to educate a well-rounded and competitive younger generation capable of adapting to contemporary demands. Young people who have received a high-quality education, possess creative and critical-thinking skills, and are proficient in information technologies can contribute to the country's sustainable development. Improving the education system is therefore one of the key priorities in Kazakhstan's efforts to achieve its goal of becoming one of the world's 30

most developed countries. Accomplishing this objective requires the modernization of educational content, the enhancement of teachers' professional competence, and the effective integration of innovative technologies into the learning process. Relevant publications also consider how chemistry and related scientific subjects can be taught to learners with diverse educational needs. Their findings contribute to the development of differentiated instruction, inclusive classroom practices, and the introduction of emerging scientific topics, such as biotechnology (Bani Ahmad, 2026; Shatayeva et al., 2022).

The primary goal of every educator is not merely to provide students with ready-made information but to enable them to search for, understand, and apply knowledge independently. Teachers should create a supportive learning environment that develops students' research abilities and encourages them to formulate questions, analyze and compare information, and draw evidence-based conclusions. In this regard, we reviewed the work of several researchers. These studies situate chemistry within the broader context of STEM education, alternative instructional formats, and the digital preparation of future science teachers. They are relevant to interdisciplinary projects, creative problem-solving, and the integration of classroom-based and off-campus learning (Eckhaus & Davidovitch, 2026; Gilmanshina, Gilmanshin, et al., 2020; Skakov et al., 2025). It is also important to enhance learners' interest in lessons, support their creative engagement, and encourage them to express their ideas freely and collaborate effectively in groups. Such a learning process promotes the development of responsibility, independence, inquisitiveness, initiative, and communication skills.

The effectiveness of an interactive learning system is closely associated with the appropriate use of the graphical, audio, video, and multimedia capabilities of information and communication technologies (ICT) in the educational process. Visual materials, animations, video lessons, virtual activities, and electronic educational resources help present complex topics to students in an accessible and engaging manner. The use of interactive games, digital exercises, assessments, and simulation software encourages students' active participation in lessons and enables them to consolidate the knowledge they have acquired. Relevant studies examine contextual, STEM-based, and mobile approaches that may increase student engagement and academic achievement in secondary chemistry education. They are particularly useful for planning introductory chemistry lessons and developing digital activities on chemical bonding for lower-secondary students (Abdikayumova & Madybekova, 2026; Dewi et al., 2022; Shagrayeva et al., 2025).

Interactive approaches transform students into active participants in the learning process. While completing learning tasks, students make decisions independently, analyze their mistakes, evaluate their results, and seek to improve their knowledge. This form of instruction facilitates the effective and meaningful acquisition of learning material, increases students' cognitive interest, and develops their information-processing skills. Therefore, the integrated use of information systems, ICT tools, and interactive methods constitutes an important prerequisite for improving the quality of education.

Within this instructional system, educators can create their own teaching materials and assessments. The development of such resources using text and graphics editing software and related digital environments – including Microsoft Word, Microsoft PowerPoint, Adobe Flash, and Adobe Photoshop – should be standardized through a unified system for producing educational content. Electronic publications, including electronic textbooks and instructional resources, are aligned with the curriculum and its individual sections and are

designed to support and optimize automated knowledge assessment. They also make it possible to determine individual learning pathways and organize various types of learning activities.

In Kazakhstan, the content of an electronic textbook is approved by the relevant government authorities, thereby granting it the status of an officially recognized instructional resource. Such textbooks enable students to study educational material at a pace appropriate to their individual needs, locate the required information quickly, and consolidate theoretical knowledge through interactive activities. Relevant studies address the pedagogical and digital competencies of preservice chemistry teachers and the relationship between university-based preparation and school practice. They provide a methodological foundation for improving lesson design, technology integration, and future teachers' professional readiness (Bekbenova et al., 2026; Karmanova et al., 2024; Meirbekov et al., 2025).

The effectiveness of electronic learning tools in the educational process has been examined in the work of numerous international scholars and substantiated both scientifically and empirically. Researchers have emphasized that multimedia materials and interactive activities enhance learners' cognitive engagement and facilitate their understanding of complex topics. Kazakhstani scholars have similarly reported that electronic learning resources positively affect the quality of education, students' interest in the subject, and the development of independent learning skills. Another group of studies focuses on experimental design, the transformation of experiments in the natural sciences, and the integration of virtual laboratories into school education. These publications can support the development of safe, visually engaging, and inquiry-based practical activities in chemistry (Berdi et al., 2025; Sagitova et al., 2020; Zhunissova et al., 2025).

In this context, the development of an integrated methodological system for computer-assisted instruction is one of the most pressing issues in improving the teaching and learning process. Such a system should ensure coherence among the content of the learning material, pedagogical methods, digital resources, and tools for assessing learning outcomes. To use an electronic resource effectively, attention should be given not only to its visual presentation but also to its alignment with learning objectives, the systematic organization of its content, and its capacity to provide feedback to students.

MozaBook is one of the contemporary digital learning environments capable of meeting these requirements. This platform integrates the capabilities of conventional and electronic instruction, allowing lesson content to be supplemented with interactive activities, video materials, animations, three-dimensional models, and virtual experiments. Examining the capabilities of MozaBook therefore facilitates a gradual transition from the theoretical foundations of electronic textbook use to their practical implementation in educational settings.

In summary, the use of electronic educational resources can enhance students' learning and cognitive engagement, improve the quality of knowledge acquisition, and increase academic achievement. In chemistry education in particular, digital technologies facilitate the visual explanation of complex concepts and processes, the simulation of experiments, and the consolidation of theoretical knowledge through practical activities. From this perspective, integrating electronic educational resources into the learning process represents an important and timely priority for the contemporary education system.

The aim of this study is to provide a theoretical and empirical justification for the effectiveness of electronic educational resources and the capabilities of the MozaBook platform in teaching the school inorganic chemistry course. The study also aims to develop an interactive electronic textbook designed to increase students' interest in chemistry, enhance their cognitive engagement, and improve the quality of their learning outcomes.

To achieve the aim of the study, the following objectives were established:

- to review research conducted by Kazakhstani and international scholars on electronic learning systems, electronic textbooks, and interactive educational technologies;
- to analyze and compare the theoretical propositions and empirical findings reported in the reviewed studies concerning the use of electronic educational resources;
- to identify the pedagogical potential of electronic educational resources in teaching inorganic chemistry at school;
- to examine the effects of electronic resources on students' learning motivation, cognitive engagement, and learning outcomes;
- to investigate the educational capabilities and tools of the MozaBook platform, as well as the specific features of its application in chemistry lessons;
- to develop a new interactive electronic textbook on the MozaBook platform in accordance with the school inorganic chemistry curriculum;
- to incorporate theoretical materials, illustrations and diagrams, videos, three-dimensional models, virtual experiments, interactive activities, and assessment questions into the electronic textbook;
- to evaluate empirically the effectiveness of using the developed electronic textbook in the educational process;
- to analyze the research findings and develop methodological recommendations for using the MozaBook electronic textbook in inorganic chemistry lessons.

As part of the study, a new electronic textbook for inorganic chemistry was developed using the MozaBook platform. The electronic textbook comprises theoretical sections, topic-specific illustrations, three-dimensional models, video and animated materials, virtual laboratory experiments, interactive exercises, consolidation activities, and tests for assessing students' knowledge. Its content is designed to present learning material in a visual, accessible, and systematically structured manner, while also supporting students' independent learning and their practical application of acquired knowledge.

## Methodology

The study was conducted among eighth-grade students at at Secondary School No. 67 in Shymkent and Secondary School No. 245 in the Zhanakorgan District. The purpose of the study was to determine the effectiveness of an electronic textbook developed using the MozaBook platform in chemistry instruction and to evaluate its effects on students' learning outcomes, academic motivation, cognitive engagement, and independent learning skills.

The methodological framework of the study comprised systemic, activity-based, learner-centered, and competency-based approaches. The study was informed by the work of Kazakhstani and international scholars on the use of electronic educational resources, chemistry teaching methodology, the principle of visualization, problem-based learning, and the gradual development of students' cognitive activity. These approaches made it possible to organize the electronic textbook's content in accordance with students' age-related characteristics, the curriculum, and subject-specific learning objectives.

The study was conducted in three stages: preparatory, experimental, and final.

During the preparatory stage, scientific and methodological literature concerning the use of electronic learning systems and digital educational resources in chemistry lessons was analyzed. The principal findings of Kazakhstani and international researchers were compared, and the pedagogical, psychological, and methodological requirements for developing electronic textbooks were identified.

In addition, the seventh-grade chemistry curriculum and its learning objectives were analyzed, and topics for inclusion in the electronic textbook were selected. Diagnostic activities, questionnaires, and assessment tasks were prepared to determine students' initial level of knowledge. Students' digital literacy skills and level of interest in chemistry were also examined during the initial assessment.

The technological and didactic capabilities of the MozaBook platform were investigated to support the digitalization of instructional materials. Methods of integrating theoretical information, illustrations, videos, animations, three-dimensional models, virtual laboratory activities, interactive exercises, and assessments within the platform were identified.

During the experimental stage, an electronic chemistry textbook was developed using the MozaBook platform and implemented in the learning process. Its content was structured according to the logical sequence of topics presented in the curriculum. The textbook incorporated theoretical materials, visual resources explaining chemical concepts and processes, videos, three-dimensional models, virtual experiments, interactive exercises, problem-based tasks, and assessment questions.

The electronic textbook was used both under the teacher's supervision during chemistry lessons and by students during independent study. Students reviewed theoretical materials, examined models of chemical phenomena, conducted virtual experiments, completed interactive activities, and assessed their knowledge through topic-based tests. The use of information and communication technologies enabled the learning materials to be presented visually and facilitated the explanation of complex chemical processes.

The study employed observation, question-and-answer activities, problem-based learning, testing, questionnaires, and comparative analysis.

The observation method was used to monitor systematically students' participation in lessons, their patterns of interaction with the electronic textbook, their independent completion of tasks, and their involvement in individual and group work. The observations recorded students' interest in the learning material, task-completion speed, questioning behavior, ability to justify their answers, and level of interaction with others.

Both oral and written forms of assessment were used to evaluate students' knowledge. Oral assessment was conducted through individual, group, and whole-class questioning. During lessons, students provided brief answers, completed tasks at the board, and worked with task cards and electronic exercises. Written assessment included topic-based tests, independent assignments, brief summative tasks, and assessment exercises incorporated into the electronic textbook.

The question-and-answer method was used to determine students' understanding of the topics and develop their ability to express ideas freely and analyze information. Students were asked clarifying, analytical, and summarizing questions concerning the content of the learning materials. They justified their responses, discussed their classmates' views, and drew conclusions regarding the issues under consideration. This method encouraged students to work with information independently, formulate questions, and systematize their acquired knowledge.

Problem-based learning was implemented by creating problem situations related to chemical phenomena and experiments. Students were presented with theoretical and practical tasks for which no predetermined answers were provided. They identified the problem, formulated hypotheses, searched for relevant information in the electronic textbook, used models and virtual experiments available on the MozaBook platform, and justified their solutions. The teacher coordinated students' activities and supported them through guiding questions.

During the final stage, the initial and final assessment results were compiled and comparatively analyzed. The quantitative analysis examined the number of completed tasks, the proportion of correct answers, and changes in students' level of knowledge acquisition. The qualitative analysis considered classroom observation data, students' responses, their approaches to completing tasks, and their opinions of the electronic textbook.

To ensure the reliability of the research findings, consistent assessment criteria were applied, and the initial and final tasks were matched in terms of difficulty. Data obtained through observation, testing, question-and-answer activities, and student feedback were compared and triangulated. This comprehensive evaluation made it possible to assess the outcomes of using the electronic textbook on the basis of multiple data sources rather than a single indicator.

## Results and Discussion

The use of the electronic textbook developed on the MozaBook platform enabled the learning material to be presented in textual, graphical, video-based, and interactive formats. The visualization of chemical phenomena through animations and three-dimensional models helped students understand processes that are difficult to observe directly. Virtual experiments allowed students to preview the stages of laboratory procedures, review safety regulations, and analyze experimental results.

While completing interactive activities and tests, students received immediate feedback on their responses and were able to correct their errors. This facilitated the review and reinforcement of the learning material and contributed to the development of self-assessment skills. The problem-based tasks were designed to encourage students not merely to memorize ready-made information but also to analyze, compare, and apply it in practical contexts.

The experimental study was conducted among seventh-grade students at Secondary School No. 67 in Shymkent and Secondary School No. 245 in the Zhanakorgan District (Figure 1).

**Figure 1**
*Photographs Taken During the Pedagogical Experiment*

During the study, the electronic textbook developed using the MozaBook platform was integrated into chemistry lessons, and its effects on students' knowledge acquisition, cognitive engagement, learning motivation, and independent learning skills were examined.

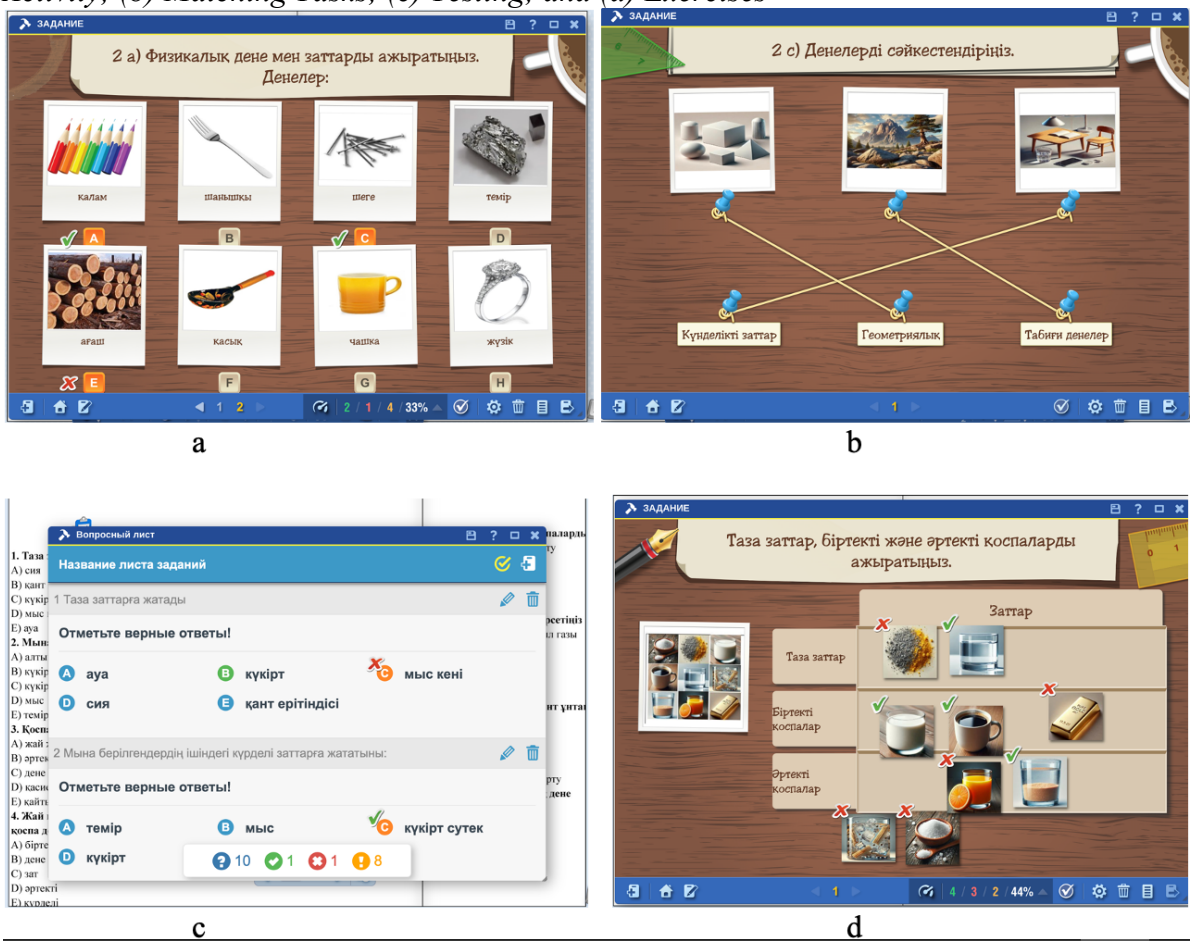
At the initial stage of the experiment, students' level of chemistry knowledge, skills in using digital resources, and interest in lessons were assessed. The diagnostic results revealed that some students experienced difficulties in explaining chemical concepts, applying theoretical knowledge in practical contexts, and identifying relationships among chemical processes. Furthermore, some students demonstrated a low level of classroom participation when working with conventional text-based materials.

During the formative stage of the experiment, chemistry lessons were organized by integrating conventional teaching methods with the capabilities of the MozaBook electronic textbook. The lessons incorporated theoretical information, images, diagrams, videos, animations, three-dimensional models, virtual experiments, and interactive activities. Students worked with the electronic textbook both independently and under the teacher's supervision. The findings of the present study are consistent with those reported by the following researchers. These studies situate chemistry within the broader context of STEM education, alternative instructional formats, and the digital preparation of future science teachers. They are relevant to interdisciplinary projects, creative problem-solving, and the integration of classroom-based and off-campus learning (Eckhaus & Davidovitch, 2026; Gilmanshina, Gilmanshin, et al., 2020; Skakov et al., 2025).

The three-dimensional models and animations available on the MozaBook platform facilitated the visual explanation of complex chemical concepts. Through visual materials, students examined the structure of substances, the mechanisms and characteristics of chemical reactions, and the stages of laboratory experiments. In this context, several studies conducted by Kazakhstani researchers should also be acknowledged. The publications in this group focus on the design of digital educational environments and reusable virtual learning objects. These approaches can assist teachers in organizing supplementary chemistry materials, visual explanations, and independent learning activities (Gilmanshina, Minnakhmetova, et al., 2020; Nurbekova et al., 2022). The comparative analysis indicated that students found abstract concepts easier to understand and developed a clearer understanding of the relationship between theoretical knowledge and practical activities (Figures 2 a–b).

**Figure 2**

*General Formats of Assessment Activities on the MozaBook Platform: (a) Classification Activity; (b) Matching Tasks; (c) Testing; and (d) Exercises*




The electronic textbook developed using the MozaBook platform incorporated interactive activities designed to consolidate students' theoretical knowledge, facilitate the comparison of concepts, and promote the practical application of acquired knowledge. The activities focused on the concepts of "physical object," "substance," "natural object," "geometric solid," and "substances used in everyday life." Students completed the activities directly on the screen and received immediate feedback on their responses.

During the virtual laboratory activities, students identified the purpose of each experiment, selected the necessary equipment and reagents, determined the sequence of procedures, and drew conclusions based on the results obtained. The opportunity to repeat virtual experiments several times helped students identify their errors and learn the correct procedures for completing the tasks. Their understanding of laboratory safety regulations and their experimental planning skills also improved. Systematic pedagogical observations indicated that students participated more actively in learning activities during lessons in which the electronic textbook was used. They demonstrated interest in completing interactive activities, asked additional questions about chemical phenomena, and attempted to justify their answers.

During the lessons, students remained focused on tasks for longer periods. When working individually, they became more active in independently searching the electronic textbook for relevant information, reviewing the material, and checking their results. During group

activities, students distributed responsibilities among themselves, exchanged information, and actively participated in formulating shared conclusions.

The use of the question-and-answer method helped students employ chemical terminology correctly, express their ideas coherently, and provide well-reasoned responses to the questions posed. At the initial stage of the experiment, most students limited themselves to brief responses; however, during subsequent lessons, their answers became more detailed and evidence-based. Rather than merely reproducing ready-made information, students learned to explain the causes of chemical phenomena, compare them, and provide relevant examples. Asking one another questions, listening to different viewpoints, and discussing responses collaboratively contributed to the development of their communication skills. While completing problem-based tasks, students developed the ability to identify problems, formulate hypotheses, locate relevant information, and draw conclusions. Using MozaBook resources, they considered several possible solutions to a problem and tested their hypotheses through virtual experiments or interactive models. During the initial lessons, students frequently sought the teacher's assistance when completing problem-based tasks. As the experiment progressed, they demonstrated greater independence by developing their own task-completion procedures and justifying their proposed solutions. Accordingly, the teacher's role shifted from providing ready-made answers to guiding students' inquiry and problem-solving activities. A comparison of the initial and final diagnostic assessment results revealed positive changes in students' mastery of the learning material (Figure 3). In the final assessment tasks, students correctly defined chemical concepts, explained chemical phenomena, and applied their theoretical knowledge in practical contexts.

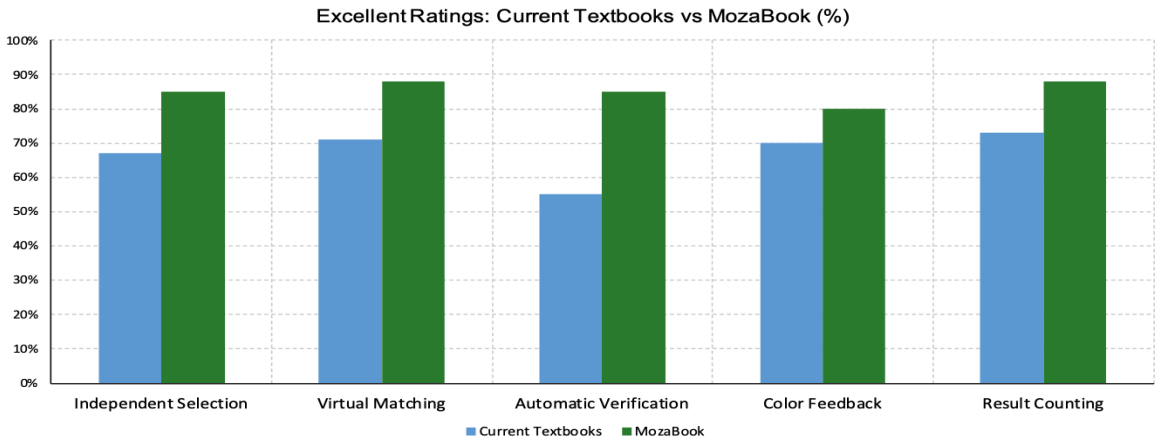
**Figure 3**

*Excellent Rating Comparison*

**Interactive Task Evaluation Results**

| Avg. Excellent Current Textbooks | Avg. Excellent MozaBook | Average Improvement | Avg. Poor Rating MozaBook |
|---|---|---|---|
| 67,2% | 85,2% | 18,0 pp | 0,0% |

**Excellent Ratings by Interactive Feature**




*Key finding: MozaBook received higher Excellent ratings for all five interactive features. The largest improvement was observed for automatic answer verification: 55% for current textbooks versus 85% for MozaBook, a gain of 30 percentage points.*

Substantial positive changes were observed in the following areas:

- searching for and selecting relevant information from electronic resources;
- correctly using chemical terminology and fundamental concepts and explaining cause-and-effect relationships in chemical processes;
- analyzing information presented through images, diagrams, and three-dimensional models;
- planning the sequence of experimental procedures and drawing conclusions based on observations;
- independently solving problem-based tasks, reviewing one's own work, and correcting identified errors.

The immediate feedback provided by the interactive tests enabled students to identify gaps in their knowledge in a timely manner. When repeating the activities, they corrected their errors and reviewed the learning material. Consequently, knowledge assessment was used not only as an evaluative tool but also as an integral component of the learning process (Table 1).

**Table 1**
*Outcomes Identified During the Pedagogical Experiment: The Pedagogical Potential of Interactive Activities and Their Effects on Students*

| Interactivity Feature | Implementation on the Platform | Pedagogical Function | Effect on the Student |
|---|---|---|---|
| Direct interaction | Selection and matching activities | Active learning | Actively participates in the task |
| Automated assessment | Identification of correct and incorrect answers | Immediate assessment | Views the result immediately |
| Color-coded feedback | Green and red indicators | Error diagnosis | Recognizes and understands errors |
| Repeated attempts | Opportunity to revise answers | Knowledge consolidation | Develops self-monitoring skills |

During the experiment, students demonstrated increased interest in chemistry. This improvement was associated with the visual presentation of the learning material, interactive activities, virtual experiments, and the opportunity to receive immediate feedback on completed tasks. For example, Table 2 presents the results of the analysis of students' learning outcomes based on their performance in the interactive activities.

Student feedback indicated that the videos, three-dimensional models, and virtual experiments included in the electronic textbook facilitated their understanding of the topics. Students regarded the opportunity to review chemical experiments and complete activities multiple times as an important advantage of the electronic textbook.

The electronic textbook also enabled differentiated activities to be assigned according to students' levels of knowledge. Students who experienced difficulties in mastering the learning material received additional visual resources and revision activities, whereas those

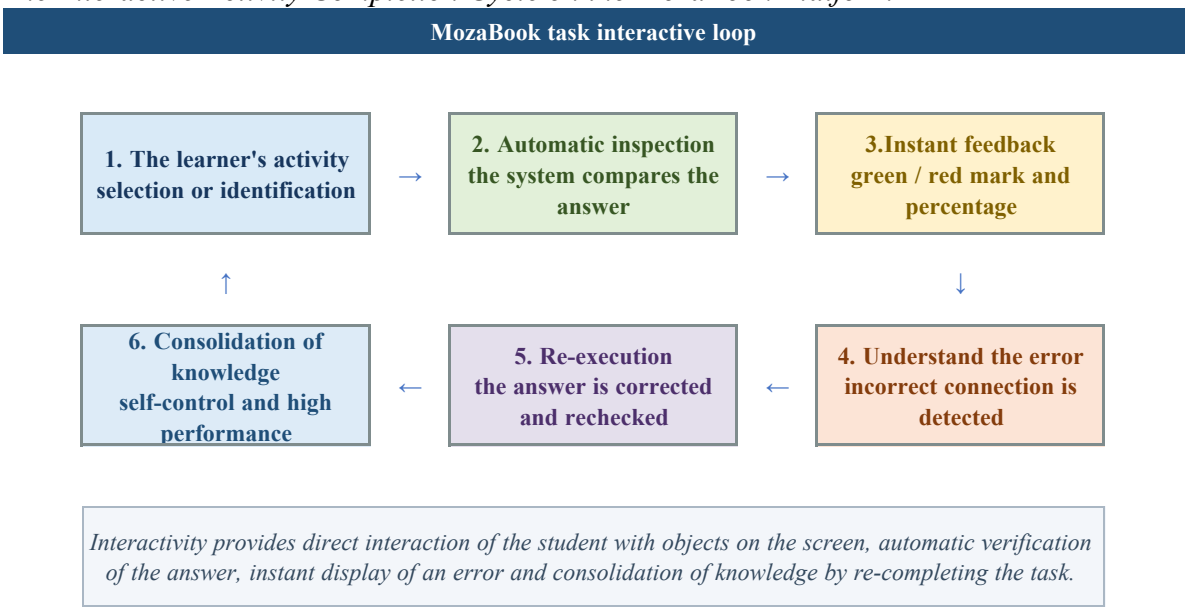
who demonstrated a strong understanding of the topics completed more advanced and problem-based tasks.

**Table 2**
*Students' Performance on Interactive Activities*

| № | Interactive Activity | Student Action | Result |
|---|---|---|---|
| 1 | Distinguishing between physical objects and substances | Selecting an object | **33%** |
| 2 | Classifying objects into the appropriate groups | Categorization | **100%** |
| 3 | Matching objects with the materials from which they are made | Matching with lines | **50%** |

The use of the MozaBook electronic textbook enabled the teacher to present learning materials systematically within a unified digital environment, organize students' learning activities, and monitor their progress in completing tasks (Figure 4). Theoretical explanation, visual demonstration, practical activities, and knowledge assessment were implemented as interconnected components of the lesson.

**Figure 4**
*The Interactive Activity Completion Cycle on the MozaBook Platform*




As shown in Figure 4, the diagram illustrates the cyclical process of completing interactive activities in MozaBook: each student action is followed by automated assessment, immediate feedback, error analysis, and a subsequent attempt. This mechanism promotes the development of self-monitoring skills, facilitates the timely correction of knowledge gaps, and supports the sustained retention of learning material.

## Conclusion

The experimental study demonstrated the effectiveness of using the electronic textbook developed on the MozaBook platform in chemistry lessons. Integrating the electronic textbook with observation, question-and-answer activities, and problem-based learning positively affected students' acquisition of the learning material, independent inquiry, completion of practical tasks, and participation in group work.

The experimental results indicated an improvement in students' level of chemistry knowledge. Students also developed a better understanding of chemical concepts and processes and strengthened their ability to apply theoretical knowledge to practical tasks. Their engagement in independent information retrieval and analysis increased, their individual and collaborative learning skills improved, and their ability to solve problems and draw conclusions developed. Overall, students demonstrated greater interest in chemistry and in using digital learning tools.

Thus, the use of the MozaBook electronic textbook not only improved students' learning outcomes but also contributed to the development of their cognitive, research, and communication skills. The findings demonstrate that the electronic textbook can be used as an effective pedagogical tool that complements conventional methods of chemistry instruction.

The study also revealed increased student interest in chemistry, greater classroom engagement, and a higher level of independent inquiry. While completing individual and group activities, students improved their ability to exchange information, justify their opinions, and reach shared decisions. The systematic use of the electronic textbook facilitated the consolidation of theoretical knowledge, supported the visual understanding of chemical concepts, and improved the quality of students' task performance.

The intended outcomes of the study were achieved: an electronic chemistry textbook was developed using the MozaBook platform, integrated into the learning process, and empirically evaluated in terms of its pedagogical potential. The applied instructional methods produced positive changes in students' knowledge acquisition, cognitive engagement, independent learning skills, and interest in lessons. The findings indicate that the MozaBook electronic textbook can serve as an effective digital educational resource that complements conventional instruction.

## Acknowledgements

This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP 22787838 "Study of interactive role digital education improving quality chemistry teachers and development Mozabook platform course [using example 7th grade]").

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that no AI or AI-assisted technologies have been used to generate, refine, or correct the content in the manuscript. The ideas, design, procedures, findings,

analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Abdikayumova, N., & Madybekova, G. (2026). Using context-based and AI-enhanced approaches to improve student engagement and achievement in secondary chemistry education. *Chemistry Teacher International, 8*(1), 37–51. https://doi.org/10.1515/cti-2025-0068

Bani Ahmad, F. (2026). Exploring the awareness and challenges faced by basic stage teachers in integrating nanotechnology into education in Jordan. *Journal of Education and e-Learning Research, 13*(2), 133–144. https://doi.org/10.20448/jeelr.v13i2.8629

Bekbenova, A., Ilyassova, G., Mukatayeva, Z., Gavronskaya, Y., Amangeldi, O., & Umirzakova, N. (2026). Pedagogical bases of ICT integration in pre-service chemistry teacher education: University-school bridge. *Frontiers in Education, 11*, Article 1729019. https://doi.org/10.3389/feduc.2026.1729019

Berdi, D., Meirbekov, A., Sarbayeva, M., Ikramov, I., & Burayeva, Z. (2025). Integrating information technologies and virtual laboratories for effective school teaching. *Journal of Chemical Education, 102*(11), 4760–4768. https://doi.org/10.1021/acs.jchemed.4c01493

Dewi, C. A., Awaliyah, N., Fitriana, N., Darmayani, S., Nasrullah, Setiawan, J., & Irwanto, I. (2022). Using Android-based e-module to improve students' digital literacy on chemical bonding. *International Journal of Interactive Mobile Technologies, 16*(22), 191–208. https://doi.org/10.3991/ijim.v16i22.34151

Eckhaus, E., & Davidovitch, N. (2026). Off-campus instruction in STEM subjects: A necessary complementary mechanism or an alternative to frontal instruction? *Education Sciences, 16*(4), Article 534. https://doi.org/10.3390/educsci16040534

Folkers, B., Castenmiller, C., Brinkman, A., & Stadermann, K. (2026). An analysis of demonstration materials in quantum physics education. *EPJ Quantum Technology, 13*(1), Article 6. https://doi.org/10.1140/epjqt/s40507-025-00460-0

Gilmanshina, S. I., Gilmanshin, I. R., Galeeva, A. I., & Abyzbekova, G. M. (2020). The digitalization of training of bachelors—the future science teachers. *IOP Conference Series: Materials Science and Engineering, 915*(1), Article 012018. https://doi.org/10.1088/1757-899X/915/1/012018

Gilmanshina, S. I., Minnakhmetova, V. A., Gilmanshin, I. R., Galeeva, A. I., & Abyzbekova, G. M. (2020). Digital educational environment: Training future teachers to design additional education. *IOP Conference Series: Materials Science and Engineering, 915*(1), Article 012017. https://doi.org/10.1088/1757-899X/915/1/012017

Grankina, V., Krasnikova, I., Smirnova, Y., Kulibaba, I., & Barinov, V. (2026). Digital transformation of the educational process in schools of Kazakhstan: The concept of intelligent assistant in school physics curriculum. *Frontiers in Education, 11*, Article 1766616. https://doi.org/10.3389/feduc.2026.1766616

Karmanova, A., Madybekova, G., Kavak, N., Ualikhanova, B., Zharylkassyn, A., & Umarova, Z. (2024). Developing the professional competence of future chemistry teachers through digital technologies: A case study of Kazakhstan. *International Journal of Information and Education Technology, 14*(8), 1119–1126. https://doi.org/10.18178/ijiet.2024.14.8.2140

Meirbekov, A., Berdi, D., Burayeva, Z., Ikramov, I., & Sarbaeva, M. (2025). A novel approach to enhancing the effectiveness of chemistry teaching by preservice teachers. *Emerging Science Journal, 9*(1), 45–71. https://doi.org/10.28991/ESJ-2025-09-01-04

Niyazova, G. Z., Duisekeyeva, B. M., Berdi, D. K., Usembayeva, I. B., & Mindetbayeva, A. A. (2026). The effects of gamified AI-supported digital learning environments on personalized learning and student engagement in school education: A systematic review and meta-analysis. *Frontiers in Education, 11*, Article 1754080. https://doi.org/10.3389/feduc.2026.1754080

Nurbekova, Z., Nurbekov, B., Maulsharif, M., Naimanova, D., & Baimendinova, A. (2022). Using virtual learning objects in educational content. *ACM International Conference Proceeding Series*, 174–178. https://doi.org/10.1145/3546118.3546138

Sagitova, R. N., Gilmanshina, S. I., Gilmanshin, I. R., Galeeva, A. I., Abyzbekova, G. M., & Saduakaskyzy, K. (2020). Digitalization of education: Transformation of a natural science experiment. *IOP Conference Series: Materials Science and Engineering, 915*(1), Article 012050. https://doi.org/10.1088/1757-899X/915/1/012050

Shagrayeva, B., Shertayeva, N., Zhorabekova, A., Baymakhanova, G., & Toktamys, N. (2025). Enhancing the quality of education through the integration of STEM technology elements in chemistry lessons. *European Journal of STEM Education, 10*(1), 1–14. https://doi.org/10.20897/ejsteme/17176

Shatayeva, A., Boranbayeva, A., Massaliyeva, Z., Batayev, D., & Makina, L. (2022). Technologies used in teaching children with special educational needs by future chemistry teacher. *World Journal on Educational Technology: Current Issues, 14*(4), 1152–1162. https://doi.org/10.18844/wjet.v14i4.7672

Skakov, M., Ramankulov, S., Nurizinova, M., Kurbanbekov, B., & Dossymov, Y. (2025). Using digital tools in STEM education and the impact on student creativity in the field of tribology. *International Journal of Evaluation and Research in Education, 14*(1), 589–597. https://doi.org/10.11591/ijere.v14i1.30220

Vidak, A., Movre Šapić, I., Mešić, V., & Gomzi, V. (2024). Augmented reality technology in teaching about physics: A systematic review of opportunities and challenges. *European Journal of Physics, 45*(2), Article 023002. https://doi.org/10.1088/1361-6404/ad0e84

Zhunissova, S., Zhussupova, L., Abyzbekova, G., & Balykbayeva, G. (2025). A review of teaching experimental design in chemistry. *Journal of Chemical Education, 102*(9), 3817–3827. https://doi.org/10.1021/acs.jchemed.5c00529

**Contact email:** science@ayu.edu.kz

# Pedagogical Technology Integration in ICT-Rich and ICT-Poor Community Learning Centres in South Africa

Khethokwakhe Goodman Chonco, Stadio Higher Education, South Africa
Khanyisile Mbatha, University of South Africa, South Africa

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

The Department of Higher Education and Training (DHET) in South Africa has made substantial investments in online learning initiatives and information and communication technology (ICT) infrastructure to enhance teaching and learning in Community Learning Centres (CLCs). However, there is limited evidence of meaningful pedagogical integration of ICT in adult education. This study examines the status of technology integration in two South African Community Learning Centres that offer adult education programmes: one with access to ICT resources and one without, both situated in semi-urban contexts. Employing a qualitative comparative case study design, data was collected through semi-structured interviews with adult learners, lecturers, and centre managers (n = 10), as well as document analysis of lesson plans, DHET ICT policy documents, annual performance plans, and staff training records. The findings indicate that both ICT-rich and ICT-poor centres exhibit minimal pedagogical integration of technology into teaching and learning practices. The availability of ICT resources did not significantly affect instructional approaches. Using Activity Theory, the study identifies systemic challenges contributing to this outcome, including the absence of DHET implementation guidelines specific to Community Education and Training (CET) colleges, a lack of institutional mandates for ICT integration by adult lecturers, insufficient pedagogical ICT training, and the absence of a dedicated learning management system (LMS) for adult learning programmes. ICT integration was primarily dependent on individual lecturer initiative rather than institutional expectations or structured support. The study contributes to adult education and ICT policy debates by demonstrating that improving technology integration in adult education requires a shift in focus from infrastructure provision to the development of CET-specific ICT policies, standard operating procedures, targeted professional development in pedagogical ICT, and a dedicated LMS for adult learning. Without these systemic supports, investments in ICT infrastructure alone are unlikely to result in meaningful pedagogical change.

*Keywords*: technology integration, adult education, Community Learning Centres, ICT policy implementation, CET colleges, pedagogical practice

iafor
The International Academic Forum
www.iafor.org

## Introduction

Information and Communication Technology (ICT) is widely viewed as an enabler of educational transformation because it can improve teaching, widen access and develop digital competencies. In South Africa, the Department of Higher Education and Training (DHET) has invested in infrastructure, online learning and national ICT policies within the Community Education and Training (CET) sector. Yet increased provision does not necessarily result in meaningful pedagogical use in Community Learning Centres (CLCs). CLCs address historical inequalities by providing disadvantaged adults with literacy, numeracy and vocational skills. Digital resources could support flexible learning, engagement and access to varied content, but integration remains uneven. Many learners also have limited prior exposure to technology, making the quality of institutional support as important as access itself. Research shows that successful integration depends on infrastructure, educator competence, institutional support and clear implementation guidance. In adult education, limited pedagogical ICT training and unclear expectations often leave technology use to individual lecturer initiative. This creates a gap between national policy intentions and classroom practice. Differences between well-resourced and poorly resourced CLCs provide an opportunity to test whether infrastructure changes pedagogy. Because equipment alone may be insufficient, it is necessary to examine how resources, policies, support and educator practices interact.

This qualitative comparative case study examines two semi-urban South African CLCs, one ICT-rich and one ICT-poor. Guided by Activity Theory, it asks how technology is integrated into teaching and learning and which systemic, institutional and pedagogical factors shape adoption.

## Literature Review

ICT can widen participation, support self-paced learning and develop employable digital competencies in adult education (Moloi & Mhlanga, 2021; Shobande & Asongu, 2022). However, infrastructure alone seldom produces sustained pedagogical change in marginalised settings (Munje & Jita, 2020; Shava, 2022). This review considers South African adult education, digital access, teaching tools and the institutional conditions required for meaningful integration.

### Adult Education Provision and Persistent Inequalities in South Africa

South African adult education developed in response to apartheid-era exclusion. Post-1994 reforms positioned ABET and CLCs as vehicles for literacy, numeracy, life skills and social participation, particularly in under-resourced communities (Meladi & Awolusi, 2020; Munje & Jita, 2020). Despite policy commitments in the White Paper on Post-School Education and Training and National Development Plan 2030, inadequate funding, infrastructure backlogs and uneven monitoring continue to shape institutional readiness for innovation (Adelabu et al., 2022; Shava, 2022).

### ICT in Education and Adult Learning: Scope and Trends

Educational ICT includes computers, mobile devices, projectors, applications and learning platforms. These tools can support interactive instruction, assessment and flexible learning for adults balancing study, work and family responsibilities (Abel et al., 2022; Mnisi et al., 2024). International experience with blended, virtual and mobile learning nevertheless identifies

recurring barriers: connectivity, device and data costs, technical support and educator preparation (Rafiq et al., 2021; Shobande & Asongu, 2022). South African adoption remains uneven. Under-resourced centres often face outdated equipment, limited connectivity and weak technical support, while socio-economic inequality reproduces digital exclusion (Aruleba & Jere, 2022; Ngcapu et al., 2022). Thus, resource availability may not translate into pedagogical integration.

### Tools, Platforms, and the Role of Learning Management Systems

Computers and mobile devices support digital content creation and access; projectors and interactive displays assist visual explanation; and learning management systems (LMSs) can organise content, communication, assessment and progress tracking (Abel et al., 2022; Meladi & Awolusi, 2020). These functions are valuable where adult learners have attendance constraints. Yet tool availability does not ensure pedagogical change. Without clear expectations, appropriate routines and institutional support, ICT use often remains occasional or administrative. Costs, outdated hardware, limited technical assistance and uncertainty about digital pedagogy further make integration dependent on lecturer initiative (Mnisi et al., 2024; Shava, 2022).

### Pedagogical Approaches and Educator Competence in Adult Learning Contexts

Blended, mobile and multimedia learning can promote flexibility, motivation and participation, while low-cost smartphone communication and formative assessment may suit constrained contexts (Abel et al., 2022; Ngcapu et al., 2022). Their value depends on alignment with adult learners' needs and learning outcomes. Educator competence is therefore decisive. Many adult educators lack confidence and training in selecting tools, designing technology-supported lessons and responding to learners with varied digital experience. Rapid technological change, low confidence and fear of failure can reinforce resistance among educators and learners (Meladi & Awolusi, 2020; Shava, 2022).

### Professional Development and Institutional Support for Integration

Professional development can improve content delivery, assessment and blended learning, but access is uneven and participation is constrained by time, funding, connectivity and data costs. Generic or once-off training rarely transfers into practice without mentoring, follow-up and institutional reinforcement (Hashemi et al., 2022; Islam & Aldaihani, 2022; Rana, 2023).

### The Digital Divide and Unequal Access in Adult Education

The digital divide encompasses unequal devices, connectivity and skills shaped by socio-economic status, geography and infrastructure. In adult education it can entrench inequality by privileging connected learners and centres (Hunduma & Seyoum, 2023; Munje & Jita, 2020). Responses include infrastructure and connectivity support, subsidised devices and data, offline solutions, educator development and partnerships (Ramnarain & Ndlovu, 2023; Shava, 2022).

### Synthesis and Gap Addressed by the Present Study

The literature identifies a persistent implementation gap: limited pedagogical integration can continue even where infrastructure exists because mandates, pedagogical training, technical

support and routine learning platforms are absent. The present comparison of an ICT-rich and ICT-poor CLC uses Activity Theory to examine how these systemic conditions shape practice.

## Methods

### Study Design and Setting

This study adopted an interpretivist qualitative comparative case study design. It enabled context-sensitive comparison of two semi-urban South African CLCs with contrasting ICT resources: one ICT-rich and one ICT-poor. The interpretivist orientation centred participants' meanings and experiences while examining how institutional and systemic conditions shaped practice (Creswell & Creswell, 2022; Sebele-Mpofu, 2020).

### Participants and Sampling

Purposive sampling recruited ten information-rich participants: four adult educators, four adult learners and two centre managers. Each CLC contributed two educators, two learners and one manager, enabling triangulation of institutional, classroom and learner perspectives (Bell et al., 2022).

Educators represented Mathematics, English/Life Orientation, sewing or skills programmes, and ICT. Learners were enrolled in ABET Level 4 and/or INCT Level 4 and had varied educational and digital backgrounds. Managers provided perspectives on resources, staff development and departmental alignment.

### Data Collection Procedures

Data were generated through semi-structured interviews and document analysis, allowing participant accounts to be compared with institutional and policy evidence (Bowen, 2009).

### Semi-structured Interviews

All ten participants completed in-person semi-structured interviews lasting approximately 45-60 minutes. With consent, interviews were audio-recorded and supported by field notes. Questions and probes explored ICT access, classroom use, professional development and enabling or constraining institutional conditions.

### Document Analysis

Document analysis covered lesson plans, DHET ICT policies, annual performance plans, staff training records and relevant public strategy documents. Documents were purposively selected and coded for infrastructure, implementation guidance, mandates, professional development and LMS provision. Comparison with interview data identified convergence and divergence between policy, practice and the two cases.

### Data Management and Analysis

Recordings were transcribed verbatim and checked against audio and field notes. Thematic analysis involved familiarisation, inductive coding, category refinement and theme development across participant groups and cases. Activity Theory then guided interpretation

of tools, rules, community, division of labour and supports such as training, technical assistance and an LMS.

### Trustworthiness

Credibility was supported through triangulation, member checking and peer debriefing. An audit trail of sampling, transcription, coding and theme development strengthened dependability; contextual description supported transferability; and reflexive memoing and evidence-based interpretation supported confirmability.

### Ethical Considerations

Institutional ethical approval preceded data collection. Participation was voluntary and based on informed consent, including the right to withdraw. Codes protected anonymity, files were securely stored, and interviews were conducted to minimise discomfort.

## Findings

Interviews and documents generated four themes. Activity Theory was used to explain how tools, rules, community, division of labour, object and outcomes shaped integration. Across both cases, pedagogical ICT use remained minimal despite different levels of infrastructure.

### Theme 1: Tools are Present but Unevenly Accessible and Unreliable

Both sites had computer laboratories, desktops, printers and projectors; the ICT-rich CLC also had laptops, smart screens and a living lab. One manager listed “25 computers ... 4 laptops, and a living lab” (Participant 1). However, nominal availability did not ensure dependable or equitable use.

Resources were concentrated in computer studies. Participants explained that they were “aimed primarily at the learners in the computer learning area and their lecturers” (Participant 4) and that computers were available to INCT Level 4 learners (Participant 5). This restricted technology-enabled learning to specialist subjects.

Booking registers, portability restrictions and workstation sharing further reduced flexible access. Connectivity and maintenance were also unreliable: faulty computers took time to repair (Participant 7), Wi-Fi slowed under heavy use (Participant 1), and one site had no internet access (Participant 4). In one case, “maintenance relies solely on our educator” (Participant 8), revealing a weak division of labour.

Third, tool functionality was weakened by infrastructural unreliability. Participants cited faulty equipment and delayed maintenance: “Some of the computers are faulty and it takes time for them to be repaired” (Participant 7). Connectivity problems were common: Wi-Fi was “slow if there’s a high usage” (Participant 1), while others described “no internet access” (Participant 4) and that “the absence of Wi-Fi connectivity significantly limits the effective integration of ICT” (Participant 5). A notable division of labour issue emerged where maintenance was informally placed on teaching staff: “Maintenance relies solely on our educator” (Participant 8). In Activity Theory terms, these constraints reduced the capacity of tools to mediate the intended object, producing outcomes characterised by intermittent access and limited instructional integration.

### Theme 2: ICT use is Largely Administrative; Classroom Integration is Minimal in Both Centres

In both centres, ICT mainly improved preparation and administration. Lecturers typed lesson plans, photocopied and prepared documentation; one said ICT was used "for lesson preparation, such as typing documentation" (Participant 3). Tools therefore mediated administrative productivity more than pedagogical change.

Classroom integration was infrequent and dependent on individual initiative. Participants reported "no clear mandate" to use technology (Participant 2), and some used it for preparation but not teaching (Participant 3). Even the ICT-rich manager described classroom use as "very limited even though the resources are available" (Participant 1), showing that infrastructure alone did not alter instruction.

Lecturers compensated with personal laptops and connectivity: "I use my own laptop and the Centre's projector" and a personal modem (Participant 9). Learners also relied mainly on cellphones at home (Participant 6). Shifting institutional responsibility to individuals produced unequal opportunities for integration.

### Theme 3: ICT is Perceived as Valuable for Adult Learning, but Benefits are not Consistently Realised

Despite limited practice, participants associated ICT with engagement, understanding, inclusion, confidence and performance, revealing desired outcomes that the activity system did not consistently produce.

Participants said textbook-only teaching bored learners, whereas simple technology increased engagement (Participant 2); videos and interactive demonstrations improved attention (Participant 9). ICT could therefore mediate learning more effectively when aligned with adult learners' need for relevant, active participation.

Visualisation was especially valuable for procedural and workplace content. A learner explained that seeing a cashier at a till connected theory to work (Participant 6), while educators said videos offered clear examples (Participant 1). Multimodal resources also supported varied literacy levels and learning preferences. However, because use was not embedded across subjects, these benefits remained intermittent.

ICT was also valued for supporting diverse learning styles and literacy levels in adult classrooms. Participants described multimodal approaches as necessary: "Learners have different learning styles… including the use of videos and images" (Participant 10), and "some learners comprehend better when they watch videos" (Participant 4). Learners and educators associated ICT-supported clarity with greater confidence and improved assessment performance: "This also assists them during assessments because they perform better" (Participant 2), and "access to technology can make a significant difference… contributing to improved academic performance" (Participant 6). However, because ICT use was not embedded across subjects and depended on individual initiative, these outcomes were reported as intermittent rather than systemic.

### Theme 4: Systemic Constraints, Rules, Capacity, and Platform Gaps Limit Pedagogical Integration

The strongest explanation for limited integration lay in weak rules, capacity and institutional supports rather than devices alone.

ICT was optional because no mandate or accountability mechanism existed. A manager confirmed that there was "no compulsory requirement" to use technology in teaching (Participant 1), leaving centres without shared routines or peer accountability.

Training focused on basic digital literacy and Word, Excel and PowerPoint rather than lesson design, learner engagement or ICT-based assessment. One lecturer reported no formal training on designing technology-supported lessons (Participant 2), while another requested training on learner management platforms (Participant 3).

Professional development was also uneven: non-ICT lecturers said funded opportunities were reserved for ICT lecturers (Participant 9). This division preserved ICT as a specialist domain rather than a shared pedagogical responsibility.

Participants requested sustained workshops, collaborative practice and institutional support. Interview and documentary evidence therefore converged: without CET-specific guidance, clear mandates, pedagogical ICT training and an LMS, technology could not routinely mediate adult learning.

## Summary of Findings Aligned to the Study Purpose

In summary, tools were unevenly accessible, regulated and unreliable; ICT mainly served administration; pedagogical use remained minimal in both centres; and perceived benefits were not systemically realised. Activity Theory locates the mismatch in weak rules, fragmented responsibilities, limited pedagogical development and absent platform support.

## Discussion

The findings explain why DHET investment did not produce meaningful pedagogical integration in either CLC. Activity Theory highlights misalignment among tools, rules, division of labour and the object of technology-enhanced adult learning.

### Infrastructure Provision: "Tools" Without Dependable Usability or Equitable Access

Infrastructure existed, especially in the ICT-rich centre, but faulty devices, connectivity problems, slow repairs and subject-based allocation weakened its usability. This supports research distinguishing access from effective use (Meladi & Awolusi, 2020; Shava, 2022) and shows that inequity also arises from how resources are distributed within an activity system. Concentrating devices in computer studies prevented whole-centre innovation.

A key contribution of this study is showing that inequity in access was not only a "resource deficit" but also a product of how tools were positioned within the activity system. ICT resources were concentrated within the computer learning area, reinforcing ICT as a specialised domain rather than a cross-cutting pedagogical resource. This aligns with concerns that subject-based allocation of ICT tends to entrench disparities in teaching quality and learner exposure,

rather than enabling whole-centre innovation (Munje & Jita, 2020; Patricia et al., 2023). From an Activity Theory perspective, the community of practice (CLC staff and learners) interacted with tools through subject-based pathways, limiting shared, cross-curricular integration.

**Regulated Access: "Rules" Designed for the Protection of Pedagogy**

Booking systems, registers and portability restrictions protected scarce assets but constrained responsive lesson design. In Activity Theory terms, local rules required pedagogy to fit access slots instead of learning needs. This contradicted broader digital policy intentions and made ICT an occasional add-on rather than a routine resource.

The results also imply a contradiction between DHET's broader digital learning intentions and the local operational rules governing access to technology. This contradiction is central to Activity Theory: policy aspirations and local practices can pull activity in different directions, leading to "stalled" or superficial implementation. The study therefore supports the abstract's claim that improving ICT integration requires moving beyond infrastructure provision to CET/CLC-specific operational procedures that prioritise pedagogical use.

**Limited Classroom Integration: The "Object" Shifts From Pedagogy to Administration**

In both centres, technology chiefly supported typing, printing and preparation. Tools therefore mediated administrative productivity rather than learner-centred pedagogy, a pattern reported in other under-resourced settings (Meladi & Awolusi, 2020; Yalley, 2022).

The ICT-rich site did not achieve stronger classroom integration than the ICT-poor site. This central finding confirms that devices are insufficient without institutional expectations, culture, training and support (Ngcapu et al., 2022; Shava, 2022).

**Pockets of Pedagogical Use: Where Tools Matched Adult Learning Needs**

The strongest pedagogical examples occurred in practical subjects, where videos and projected demonstrations supported procedures and workplace relevance. These uses reduced abstraction and aligned with adult learners' experience, but remained localised innovations dependent on individual lecturers rather than centre-wide practice (Adelabu et al., 2022; Kilag et al., 2023).

However, the study shows these practices were not institutionalised and remained dependent on individual lecturer initiative. In Activity Theory terms, the activity system did not redistribute innovation across the community; instead, innovations remained localised, reinforcing uneven outcomes across subjects. This supports the abstract's conclusion that ICT integration was shaped more by lecturer agency than by structured institutional expectations.

**Perceived Pedagogical Value: Positive "Outcomes" Imagined, but Not Systemically Produced**

Participants recognised ICT's potential to improve engagement, comprehension, inclusion and confidence, yet these outcomes were inconsistent because technology was not embedded in routine teaching. The activity system thus remained at an early stage: stakeholders valued ICT, but enabling conditions were absent.

## Professional Development and Institutional Support: A Weak "Division of Labour" and Missing Platform Mediation

Professional development was basic, uneven and operational rather than pedagogical, leaving non-ICT educators under-supported. Technical training without attention to teaching design is unlikely to transform practice (Shava, 2022). Activity Theory identifies a division-of-labour failure: responsibility moved to educators and their personal devices, while institutional training, technical support and accountability remained weak. The lack of an LMS also removed a stable structure for digital content, assessment and communication.

## Implications: Shifting the Focus From Infrastructure to System Design

Meaningful integration therefore requires systemic alignment: usable tools; enabling rules and implementation guidance; equitable training and technical support; leadership accountability; and a low-bandwidth LMS. The study's contribution is to show that the key question is not whether devices exist, but whether the CET activity system makes pedagogical use routine, equitable and sustainable.

## Conclusion and Recommendations

Recommendations respond to four findings: infrastructure was uneven and unreliable; ICT use was mainly administrative; recognised pedagogical benefits were inconsistently realised; and training was basic and unequal. Priority should therefore shift from equipment alone to CET-specific implementation, capability and institutional routines.

### Policy-Level Recommendations (DHET/CET System)

DHET should translate national digital goals into a CET/CLC-specific policy with minimum classroom expectations, role clarity, timelines, monitoring, standard operating procedures and sample ICT-integrated lesson and assessment templates. A minimum service package should specify reliable or offline connectivity, maintenance targets, presentation capability, secure storage, user support and equitable access across learning areas. Need-based funding should account for connectivity, resource poverty and maintenance capacity, with protected budgets for internet access, equipment renewal and technical support. Support affordable access strategies for adult learning, Affordable access should include negotiated data, community Wi-Fi, zero-rated educational platforms and offline repositories suited to adult learners. System monitoring should track pedagogical indicators, including ICT-supported lessons across subjects, staff development participation and LMS use, rather than devices delivered alone.

### Institutional Recommendations (CLC and CET College Management)

Each centre should maintain a concise implementation plan covering access, cross-subject expectations, training, support responsibilities and term targets. Access rules should balance asset protection with predictable learner laboratory time, educator preparation and flexible use by non-ICT subjects. CLCs need fault-reporting, escalation and repair protocols. Where a technician is unaffordable, CET colleges could share roaming support across clustered centres. Institutions should provide pooled staff devices and connectivity so integration does not depend on personal resources. A mobile-compatible, low-bandwidth LMS should support announcements, downloadable materials, simple assessments and feedback, normalising blended routines for adults with irregular attendance.

## Professional Development Recommendations (Pedagogical ICT Capability)

Professional development should move from generic digital literacy to lesson design, learner engagement, assessment and classroom management with technology. Opportunities must include educators across subjects and age groups, using paced learning, peer mentoring and non-punitive support. Continuous professional learning communities, coaching, shared lesson plans and observation should replace once-off workshops. Embed ICT integration into developmental performance support Developmental appraisal should include supported ICT goals, coaching, recognition and access to resources rather than punitive compliance.

## Pedagogical Recommendations (Classroom and Curriculum Practice)

Classroom practice should prioritise visualisation, workplace relevance and structured practice, extending successful uses in practical subjects to literacy, numeracy and life skills. Adopt low-tech, high-impact integration pathways. Low-tech, high-impact pathways may use projectors, offline videos, local repositories, mobile activities and brief formative assessments. Learners need orientation in device use, online safety and LMS access so technology does not create new exclusions. Realistic minimum targets, supported by shared resources and co-planning, can move integration from ad hoc practice to routine use.

## Recommendations for Future Research

Future research should include additional provinces and CLCs and use longitudinal or mixed methods to examine whether policy guidance, LMS adoption, and pedagogical training improve attendance, retention, completion, and assessment. Further study should address older adults' experiences and compare digital transformation across CET, TVET and schools.

## References

Abel, V. R., Tondeur, J., & Sang, G. (2022). Adult educator perceptions about ICT integration into classroom instruction. *Education Sciences, 12*(9), 609. https://doi.org/10.3390/educsci12090609

Adelabu, F. M., Alex, J. K., Ngwabe, A., Tatira, B., & Boateng, S. (2022). Creation of innovative teaching spaces with Gamma Tutor: A techno blended model for rural mathematics teaching. *Journal of Education and e-Learning Research, 9*(4), 249–257. https://doi.org/10.20448/jeelr.v9i4.4244

Aruleba, K., & Jere, N. (2022). Exploring digital transforming challenges in rural areas of South Africa through a systematic review of empirical studies. *Scientific African, 16*, e01190. https://doi.org/10.1016/j.sciaf.2022.e01190

Bell, E., Bryman, A., & Harley, B. (2022). *Business research methods* (6th ed.). Oxford University Press.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Creswell, J. W., & Creswell, J. D. (2022). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE Publications.

Hashemi, A., Si Na, K., Noori, A. Q., & Orfan, S. N. (2022). Gender differences on the acceptance and barriers of ICT use in English language learning: Students' perspectives. *Cogent Arts & Humanities, 9*(1), 2085381. https://doi.org/10.1080/23311983.2022.2085381

Hunduma, C. M., & Mekuria, Y. S. (2023). Constraints hindering ICT integration in Ethiopian public secondary schools: A literature review. *International Journal of Membrane Science and Technology, 10*(2), 691–702. https://doi.org/10.15379/ijmst.v10i2.1271

Islam, M. A., & Aldaihani, F. M. F. (2022). Justification for adopting qualitative research method, research approaches, sampling strategy, sample size, interview method, saturation, and data analysis. *Journal of International Business and Management, 5*(1), 1–11. https://doi.org/10.37227/JIBM-2021-09-1494

Kilag, O. K. T., Segarra, G. B., De Gracia, A. M. L., Del Socorro, A. S., Abendan, C. F. K., Camangyan, G. A., & Mahasol, E. T. (2023). ICT application in teaching and learning. *Science and Education, 4*(2), 854–865.

Meladi, O., & Awolusi, O. (2020). The influence of information communication technology (ICT) integration on teaching and learning in South African schools. *Journal of Education and Vocational Research, 10*, 47–64. https://doi.org/10.22610/jevr.v10i2(V).3023

Mnisi, B., Mtshali, T., & Makgato, M. (2024). Moving beyond the challenges of learning through technologies: The current status of ICT integration in South African schools. *Journal of Education and e-Learning Research, 11*(1), 128–134. https://doi.org/10.20448/jeelr.v11i1.5396

Moloi, T., & Mhlanga, D. (2021). Key features of the fourth industrial revolution in South Africa's basic education system. *Journal of Management Information and Decision Sciences, 24*(5), 1–20.

Munje, P. N., & Jita, T. (2020). The impact of the lack of ICT resources on teaching and learning in selected South African primary schools. *International Journal of Learning, Teaching and Educational Research, 19*(7), 263–279. https://doi.org/10.26803/ijlter.19.7.15

Ngcapu, S. R., Simelane-Mnisi, S., & Mji, A. (2022). Strategy as plan for technology integration to reposition lecturers for the new normal in higher education. In *New updates in e-learning*. IntechOpen. https://doi.org/10.5772/intechopen.102078

Patricia, M. M., Isaac, M. T., & Manto, R. S. (2023). Enhanced creativity with Mini-PAT: A case of grade 9 technology educators in Sekhukhune East. *African Perspectives of Research in Teaching and Learning, 7*(1), 110–121.

Ramnarain, U., & Ndlovu, M. (2023). *Information and communications technology in STEM education: An African perspective*. Routledge.

Rana, K. (2023). Rural educators' perceived affordances of ICT training. *Interactive Learning Environments*, 1–15. https://doi.org/10.1080/10494820.2023.2216739

Sebele-Mpofu, F. Y. (2020). Saturation controversy in qualitative research: Complexities and underlying assumptions. A literature review. *Cogent Social Sciences, 6*(1), Article 1838706. https://doi.org/10.1080/23311886.2020.1838706

Shava, E. (2022). Reinforcing the role of ICT in enhancing teaching and learning post-COVID-19 in tertiary institutions in South Africa. *Journal of Culture and Values in Education, 5*(1), 78–91. https://doi.org/10.46303/jcve.2022.7

Shobande, O. A., & Asongu, S. A. (2022). The critical role of education and ICT in promoting sustainability in Africa. *Technological Forecasting and Social Change, 176*, 121480. https://doi.org/10.1016/j.techfore.2022.121480

Yalley, C. E. (2022). Challenges affecting the use of ICT in school administration. *Cogent Education, 9*(1), 2062893. https://doi.org/10.1080/2331186X.2022.2062893

**Contact email:** kchonco@yahoo.com

# The AI Academic Integrity Toolkit: Designing Ethical and Reflective Assessment Practices in the Age of AI

Rahme Sadikoglu, Anglia Ruskin University London, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This paper presents the AI Academic Integrity Toolkit, an open educational resource designed to support lecturers in embedding ethical and reflective AI use within assessment design and feedback. Building on the AI-Enabled Academic Integrity (AIAI) Framework (Visin, 2025), the Toolkit translates its four pillars, Ethical Access, Critical Competence, Reflexive Transparency and Educational Value, into assessment-oriented design principles. The Toolkit was developed through reflective engagement with academic integrity review processes, review of institutional policy guidance and ongoing professional discussions with educators. The Toolkit provides adaptable assessment models, suggested assessment brief language, reflection prompts and feedback resources aligned with principles of constructive alignment (Biggs & Tang, 2011) and an educative understanding of academic integrity (Bretag, 2013). Designed for potential application across undergraduate and postgraduate contexts, it aims to clarify boundaries around responsible AI use, promote student reflection and support educators in evaluating AI-assisted work. The Toolkit is presented as a design-based contribution intended to support pedagogical implementation and future evaluation, rather than as a report of empirical outcomes. By reframing integrity as an educational rather than a compliance-based practice, the Toolkit contributes an adaptable, theoretically grounded approach to integrating AI literacy and ethical reflection within higher-education assessment practices.

*Keywords:* academic integrity, AI literacy, assessment design, higher education, reflective pedagogy

iafor
The International Academic Forum
www.iafor.org

## Introduction

Generative artificial intelligence (GenAI) has unsettled established assumptions about the relationship between authorship, originality and learning in higher education. Students can now use AI systems to generate ideas, organise arguments, produce written content and revise existing work, making it increasingly difficult to regard the final submission as a straightforward representation of individual cognitive effort. These developments do not render authorship or originality irrelevant. They do, however, complicate the boundaries between legitimate assistance, collaborative knowledge production and inappropriate outsourcing, and raise important questions about what assessment is intended to evidence in AI-mediated learning environments (Cotton et al., 2024; Fawns, 2019; Kofinas et al., 2025).

Understandably, a prominent strand of the institutional and academic response has focused on academic misconduct, policy revision and the detection of AI-generated text (Cotton et al., 2024; Kofinas et al., 2025). Such responses remain necessary where AI is used deceptively or in ways that prevent students from demonstrating the required learning outcomes. However, detection alone cannot establish how a student used AI, whether its outputs were questioned, what decisions were made during the process or whether the interaction contributed to learning. Academic integrity should therefore not be reduced to the absence of misconduct. As Bretag (2013) argues, integrity also needs to be understood as an educative and developmental concern embedded within curriculum and institutional culture.

The educational challenge is consequently not only to regulate AI use, but to create assessment conditions in which students can learn to use it responsibly. The AI-Enabled Academic Integrity (AIAI) Framework developed by Visin (2025) addresses this challenge through four interconnected pillars: Ethical Access, Critical Competence, Reflexive Transparency and Educational Value. Together, these principles shift attention from the binary question of whether AI was used towards the quality, purpose and transparency of that use. Yet conceptual principles do not automatically translate into everyday assessment practice. What remains less developed is how lecturers can embed such expectations within assessment briefs, reflection activities, marking criteria and feedback without requiring extensive changes to existing modules.

This paper responds to that practical gap by presenting the AI Academic Integrity Toolkit (Sadikoglu, 2025). Building on the AIAI Framework (Visin, 2025), the Toolkit operationalises its four pillars through adaptable resources for assessment design, student reflection and feedback. Informed by constructive alignment, it proceeds from the position that ethical and critical AI use must be made visible within the assessment process if it is to become a meaningful part of student learning (Biggs & Tang, 2011). I argue that the Toolkit offers a practical route for moving academic integrity from a predominantly reactive concern towards a more educative practice, while retaining the need for clear institutional boundaries. Before presenting the Toolkit in detail, the following section situates this contribution within debates on academic integrity, GenAI and constructive alignment. The paper then explains the practice-informed design process, introduces the Additive, Embedded and Developmental integration models, and considers their intended applications and limitations.

## From Detection to Educational Integrity: Literature and Conceptual Foundations

The International Center for Academic Integrity (2021) articulates academic integrity through six fundamental values: honesty, trust, fairness, respect, responsibility and courage. These

values remain fundamental to higher education, but their institutional application has often centred on rule compliance, individual authorship and the prevention of misconduct. Bretag (2013) argues that institutions should move beyond deterrence, detection and punishment by embedding academic integrity within assessment, curriculum and institutional culture. This educative position is particularly important in the context of generative artificial intelligence, where the distinction between legitimate academic support and inappropriate substitution of student effort has become increasingly difficult to establish through the final submitted product alone.

GenAI systems can support students in generating ideas, structuring arguments, summarising information and producing written content. Their presence within educational practice therefore makes it more difficult to infer a student's individual contribution and learning from the submitted text alone (Cotton et al., 2024; Rudolph et al., 2023). This does not mean that individual responsibility or authentic learning have ceased to matter. Rather, it means that authorship is increasingly shaped by interactions between the student, the assessment task and the technological tools available to them. Fawns's (2019) account of postdigital education is useful here because it rejects the idea that digital technologies function merely as external additions to otherwise unchanged educational practices. Instead, technologies form part of the conditions through which learning, writing and knowledge production take place. The integration of AI should not, however, be assumed to produce educational value in itself. O'Dea and O'Dea (2023) observe that evidence concerning the pedagogical impact of AI in higher education remains limited, despite widespread claims about its potential. This reinforces the need to examine what a particular use contributes to learning rather than equating technological adoption with educational benefit.

These developments expose the limitations of responses that concentrate primarily on detecting whether AI has been used. Detection and regulation remain necessary where students conceal AI use, breach clearly stated assessment requirements or submit work that does not demonstrate the intended learning outcomes. However, the identification of possible AI-generated content cannot, by itself, explain how a student worked with AI, why particular outputs were accepted or rejected, or whether the interaction supported understanding. Kofinas et al. (2025) found that markers generally could not distinguish assessments with GenAI input from those without it and that authentic assessment alone did not safeguard academic integrity. The educational question is consequently broader than whether AI was present within the process. It also concerns the quality of the student's judgement, the transparency of their decisions and the relationship between AI use and the purpose of the assessment.

I therefore approach academic integrity not only as a regulatory responsibility but also as an educational practice. From this perspective, students should be supported in developing the capacity to make informed decisions about when AI use is appropriate, to question the reliability and limitations of generated outputs, and to explain how those outputs were incorporated into their work. Integrity is not replaced by AI literacy within this formulation. Rather, critical and ethical AI literacy becomes one of the contemporary conditions through which integrity is enacted. This position retains the importance of institutional boundaries while recognising that students require more than prohibitions if they are to navigate AI-mediated learning responsibly.

Moving beyond a final-product focus requires these expectations to be made visible before and during assessment. This is particularly important because assessment influences what students attend to, how they distribute their effort and how they approach learning (Gibbs & Simpson,

2005). Constructive alignment provides a basis for connecting these expectations with intended learning outcomes. Biggs and Tang (2011) argue that learning activities and assessment tasks should be aligned with those outcomes. If ethical judgement, critical evaluation and transparency are valued educational outcomes, they must therefore be made visible within assessment design rather than assumed or addressed only after suspected misconduct has occurred. Assessment briefs communicate what students are expected to do, marking criteria indicate what will be valued, and feedback helps students understand the quality of their current performance and how it might be developed (Carless & Boud, 2018; Price et al., 2012). In the context of GenAI, constructive alignment means connecting guidance on AI use with the purpose of the task, the required evidence of learning and the feedback students receive.

The AI-Enabled Academic Integrity Framework (Visin, 2025) offers a conceptual structure for this alignment. Its first pillar, Ethical Access, concerns the student's judgement about when and why AI use is appropriate within a particular educational context. Critical Competence refers to the capacity to construct purposeful prompts and evaluate AI outputs for accuracy, bias, relevance and disciplinary suitability. Reflexive Transparency extends beyond a basic declaration by asking students to explain which tools were used, how outputs were selected or modified, and why particular decisions were taken. Finally, Educational Value considers whether AI contributed meaningfully to understanding or skill development rather than functioning simply as a shortcut to task completion. The four pillars are interconnected: appropriate access without critical evaluation is insufficient, while transparency has limited educational value if it does not reveal purposeful engagement with learning.

An established practical response to these challenges is the Artificial Intelligence Assessment Scale (AIAS), which enables educators to specify and communicate the permitted role of GenAI within an assessment and has subsequently been developed as a framework for task redesign and assessment validity (Perkins, Furze, et al., 2024; Perkins, Roe, et al., 2025). The AI Academic Integrity Toolkit is complementary to this work rather than a substitute for it. Whereas the AIAS establishes the role and scope of permitted AI use within a task, the Toolkit focuses on how students explain and evaluate their decisions, how that evidence can inform assessment criteria, and how feedback can support further development.

The AIAI Framework consequently offers a basis for moving from the binary distinction between AI use and non-use towards a more considered evaluation of purpose, process and judgement. What remains less developed, however, is how these principles can be incorporated into the everyday structures of assessment without requiring complete curriculum redesign. The AI Academic Integrity Toolkit responds to this practical gap by operationalising the four pillars through adaptable assessment models, student reflection and feedback practices (Sadikoglu, 2025). The following section explains the practice-informed design process through which these resources were developed.

## Developing the AI Academic Integrity Toolkit: A Practice-Informed Design Approach

As established in the previous section, the central difficulty is not simply the absence of ethical principles, but the lack of practical routes through which they can be embedded within assessment. The AI Academic Integrity Toolkit was developed as a practice-informed response to this problem (Sadikoglu, 2025). The Toolkit was therefore not conceived as a separate policy document, nor as a replacement for existing academic integrity procedures. Instead, it was designed to help educators incorporate ethical and critical AI use into the ordinary structures through which assessment expectations are communicated and learning is evaluated.

Its development was informed by four interconnected sources. First, reflective engagement with academic integrity review processes drew attention to an evidential gap that arises when attention is directed primarily towards the final submission. Identifying which tools were used does not necessarily reveal why particular outputs were accepted, rejected or modified, or how those decisions related to the intended learning outcomes. Second, institutional guidance provided an important regulatory context by clarifying the distinction between permitted educational support and forms of AI use that may prevent authentic demonstration of learning. Third, scholarship on constructive alignment, assessment literacy, feedback and academic integrity helped position these practical concerns within a wider pedagogical framework (Biggs & Tang, 2011; Bretag, 2013; Carless & Boud, 2018; Price et al., 2012). Finally, ongoing discussions with educators and reflection on my wider academic support practice helped shape the resource by drawing attention to disciplinary variation, validated assessment structures, staff workload and the need for accessible, discipline-sensitive student-facing language supported by concrete examples.

I describe this as a practice-informed design process rather than an empirical study. The professional discussions that informed the Toolkit were not used as formally collected research data, and no claims are made here about educator or student outcomes. No student case records, identifiable case material or formally collected case data were analysed for this paper. The role of these professional discussions was developmental: they helped identify practical questions that the Toolkit needed to address and informed decisions about its structure, flexibility and language.

Several design principles followed from this process. The Toolkit needed to be adaptable across disciplines, proportionate in relation to staff workload and sufficiently flexible to operate within different institutional and assessment contexts. It also needed to retain a developmental orientation. Students should not be expected merely to declare whether AI was used; they should be supported in explaining the reasoning behind that decision, how any AI-generated outputs were evaluated and what educational value, if any, AI use produced. For this reason, the Toolkit connects assessment design, reflection and feedback rather than approaching them as separate activities. Its purpose is not to impose a single institutional model, but to provide educators with a set of adaptable pathways through which the four AIAI pillars can be made visible in assessment practice.

## The AI Academic Integrity Toolkit: Structure and Assessment Applications

This section turns to the Toolkit itself and examines how the four pillars of Visin's (2025) AIAI Framework are translated into practical resources for assessment design, student reflection and feedback (Sadikoglu, 2025). The Toolkit is organised around the principle that expectations concerning AI use should be communicated before students complete an assessment, considered during the learning process and revisited through feedback afterwards. In this sense, academic integrity is positioned not as a separate intervention that begins only when misconduct is suspected, but as part of the ordinary design of assessment.

The Toolkit contains three connected areas of practice. The first concerns the design and communication of assessment. This includes three integration models through which lecturers may incorporate the AIAI principles into existing assessment structures, together with suggested language for assessment briefs. The second concerns student reflection. A self-assessment form invites students to explain whether AI was used and the reasoning behind that decision. Where AI was used, students explain how they constructed prompts, evaluated the

outputs produced and decided what to accept, reject or modify. The form also asks them to identify where AI contributed to the assessment and to consider whether its use supported their learning. The third concerns feedback and feedforward. The Toolkit provides traffic-light feedback, short pillar-based comments and reflective questions intended to help students identify where their engagement was strong and where further development may be needed (Sadikoglu, 2025).

These components are intended to operate together. Assessment briefs communicate what is expected, reflection makes the student's process more visible, and feedback provides an opportunity to develop future judgement. This sequence reflects the principles of constructive alignment because it connects what students are asked to do with what educators subsequently evaluate and discuss with them (Biggs & Tang, 2011). It also draws on assessment literacy, which emphasises that students need access to the language, standards and expectations through which academic work is judged (Price et al., 2012). In the context of GenAI, this means that students should not be expected to infer the boundaries of acceptable use or understand automatically what critical and transparent AI use involves.

Table 1 summarises the three integration models, which are discussed in more detail below. Together, they provide different routes for embedding the AIAI principles within assessment practice.

**Table 1**
*Three Models for Integrating AIAI Into Assessment Practice*

| Model | Approach | Potentially suitable for |
|---|---|---|
| **Additive** | Introduces AI use and academic integrity as a separate criterion within the marking scheme, allowing the four AIAI pillars to be assessed directly. | Assessments in which AI use is an explicit requirement and there is sufficient flexibility to add or revise a marking criterion. |
| **Embedded** | Incorporates relevant AIAI principles within existing assessment or marking criteria, with evidence from the self-assessment form informing how those criteria are applied. | Modules with validated or relatively fixed marking schemes, or where the AIAI principles align naturally with existing criteria. |
| **Developmental** | Uses the AIAI self-assessment as a non-graded reflective activity, supported by formative feedback on the four pillars. | Foundation, early-stage or skills-development modules where the priority is developing awareness and reflective practice. |

*Note.* Adapted from *A Practical Toolkit for Ethical and Reflective Assessment Design in the Age of AI* (Sadikoglu, 2025).

The Additive model introduces AI use and academic integrity as a distinct assessment or marking criterion. For example, a module may include a dedicated criterion that evaluates how appropriately, critically and transparently a student used AI, the decisions they made during the process and whether that use supported the intended learning outcomes. This approach makes expectations highly visible and allows the four AIAI pillars to be assessed directly rather than remaining implicit within broader criteria. It may therefore be particularly appropriate where ethical and reflective AI use forms an explicit part of the assessment. Introducing a separate criterion may, however, require formal approval or revision of existing marking

arrangements, depending on institutional and programme-level processes. The wording and weighting of the criterion should also remain proportionate to the wider assessment task so that AI use does not overshadow the disciplinary knowledge and skills being assessed. For this reason, the Additive model is most appropriate where AI use is an explicit assessment requirement rather than simply an optional form of support. The assessment brief should clearly specify the permitted scope of AI use, the evidence students are expected to provide and how their use will be evaluated. Where AI use remains optional, the Embedded or Developmental model may offer a fairer and more appropriate route.

The Embedded model incorporates the AIAI principles within criteria that already exist. A criterion concerning critical analysis, for instance, may include consideration of how a student evaluated the accuracy, relevance or bias of AI-generated material, while a criterion concerning reflective practice may include their explanation of why AI was used, which outputs were accepted or rejected and how those outputs were revised. Evidence for these judgements can be provided through the AIAI self-assessment form, submitted alongside the main assessment. The marker can then consider the student's responses when applying the relevant existing criteria, rather than awarding a separate mark for AI use. Where AI use is optional, choosing not to use it should not in itself reduce the mark; the existing criterion should continue to assess the academic capability it was designed to measure. This model may be more appropriate where adding a new criterion is impractical or where the AIAI principles align naturally with existing learning outcomes and marking criteria. Its advantage lies in connecting ethical and critical AI use with the academic capabilities that the assessment already values. At the same time, there is a risk that expectations may remain too implicit. The assessment brief should therefore identify which existing criteria relate to AI use, explain how the self-assessment form will inform marking and clarify what students are expected to disclose and reflect upon.

The Developmental model adds a non-graded reflective component to the assessment process. Students complete the AIAI self-assessment and receive formative feedback on the four pillars, without their responses contributing directly to the final mark. Unlike the Additive and Embedded models, the self-assessment is not used as evidence when applying marking criteria. This model may be particularly appropriate for foundation, early-stage or skills-development modules where the immediate priority is to build confidence and reflective habits. It is also intended to create conditions in which students can discuss their AI use more openly, because disclosure is framed as an opportunity for development rather than as evidence of misconduct. Nevertheless, reflection should not be assumed to be straightforward. Students may require examples, simplified language and formative guidance to move beyond brief declarations towards meaningful explanation. A developmental approach does not, however, override institutional academic integrity procedures. A self-assessment may indicate AI use that falls outside the boundaries stated in the assessment brief, and the Toolkit cannot replace the academic integrity procedures that may apply in such circumstances. Its role is instead to make expectations clear from the outset by stating what forms of AI use are permitted, explaining the purpose of the self-assessment and positioning accurate disclosure as part of responsible academic practice. This clarity is important for sustaining trust while avoiding the suggestion that the developmental model operates separately from institutional requirements.

Across the three models, the AIAI self-assessment provides a common structure for making students' decisions more visible, although its function differs according to the model selected. In the Additive and Embedded models, it may provide evidence that informs the application of the relevant criterion or criteria; in the Developmental model, it is used only for reflection and formative feedback. Students are asked to explain whether AI was used and the reasoning

behind that decision, describe how prompts and generated outputs were evaluated, identify the tools used and where they contributed to the assessment, and reflect on what, if anything, AI use contributed to their learning. The form therefore moves beyond simple disclosure by making the student's decisions available for reflection, assessment and feedback, as appropriate to the model being used (Sadikoglu, 2025). Table 2 provides an illustrative Embedded-model application, showing how assessment brief language and the self-assessment form can work together.

**Table 2**
*Illustrative Embedded-Model Application*

| Toolkit Element | Illustrative wording |
|---|---|
| Assessment brief | You may use generative AI tools to generate initial ideas, clarify concepts or obtain feedback on a draft. You may not submit AI-generated text as your own writing or use AI in ways that replace the learning outcomes being assessed. You must submit the AIAI self-assessment form alongside your work. Your responses will inform how the existing criteria for critical analysis and reflective practice are applied; they will not receive a separate mark. Choosing not to use AI will not disadvantage you, provided that decision is explained. |
| Self-assessment prompts | Ethical Access: Did you use AI in preparing this assessment? Explain the reasoning behind your decision in relation to the purpose of the task. Critical Competence: Where AI was used, how did you evaluate its outputs? What did you accept, reject or modify, and why? Reflexive Transparency: Which tools did you use, and where did they contribute to the submitted work? Educational Value: What, if anything, did AI use contribute to your understanding or skill development? |

*Note.* Adapted from *A Practical Toolkit for Ethical and Reflective Assessment Design in the Age of AI* (Sadikoglu, 2025).

The feedback resources follow the same pillar-based structure. Traffic-light feedback offers a concise indication of whether the student's reflection demonstrates each pillar clearly, partially or not yet sufficiently. Under Critical Competence, for example, green indicates a clear evaluation of the strengths and limitations of specific outputs; amber indicates some evaluation but limited depth; and red indicates that outputs were accepted without critical examination. Short reflective comments allow lecturers to identify specific strengths and areas for development, while feedforward questions encourage students to consider how their future decisions and practices might be improved. This future-oriented emphasis is important because feedback does not automatically influence later practice; students need guidance that is sufficiently specific and actionable to inform subsequent decisions (Carless & Boud, 2018; Gibbs & Simpson, 2005). The Toolkit does not assume that every lecturer will use all three feedback formats. Rather, the resources are intended to be selected and adapted according to module purpose, student level and staff capacity.

Taken together, these components show that the Toolkit offers three adaptable pathways rather than a single fixed model. Legitimate uses of AI vary across disciplines, while institutional policies and assessment regulations also differ. Students may also vary in their familiarity with reflective writing and their ability to explain their decisions. It is important, therefore, that the four AIAI pillars remain consistent while their practical application can be adapted to different contexts. The purpose is not to standardise AI use across disciplines or institutions, but to

provide a structure through which expectations can be made clearer and students' decisions can become a more visible part of the assessment process.

## Discussion, Limitations and Future Evaluation

The AI Academic Integrity Toolkit offers a set of adaptable pathways for translating ethical principles into assessment design. Its contribution lies not in proposing that AI use should always be encouraged, but in providing structures through which its appropriateness, purpose and educational value can be considered more explicitly. This distinction is important. The presence of AI within an assessment does not, by itself, demonstrate either meaningful learning or academic misconduct. What requires attention is how students use the technology, how their decisions relate to the learning outcomes and whether the assessment continues to provide credible evidence of their knowledge and judgement (Cotton et al., 2024; Kofinas et al., 2025). The Toolkit may therefore support a more educational approach to academic integrity by connecting expectations, reflection and feedback. By aligning expectations concerning AI use with assessment activities, criteria and feedback, it draws together the pedagogical principles discussed earlier in the paper (Biggs & Tang, 2011; Carless & Boud, 2018; Price et al., 2012). This complementarity also has a practical implication. Where an institution already uses the AIAS or another permission-based framework, the Toolkit can operate within the permitted level of use by structuring the reflection, evidence and feedback that follow (Perkins et al., 2025). AI use is consequently positioned as something that can be examined and developed throughout the assessment process rather than considered only after a possible breach has occurred.

Adaptability should not be confused with universal applicability. Disciplines differ considerably in what constitutes legitimate AI assistance, and institutional policies may restrict how assessment briefs or marking criteria can be modified. The Additive model is most appropriate where AI use forms an explicit assessment requirement and therefore depends on clear task instructions, equitable access to the relevant technology and proportionate weighting within the marking scheme. Where AI use is required, institutions should provide access to an approved tool or to tools offering sufficiently comparable functions. Assessment design should not advantage students who can purchase enhanced versions, and guidance should distinguish permitted language or accessibility support from uses that replace the learning being assessed. In the Embedded model, the relationship between the AIAI principles and existing criteria may remain insufficiently visible unless it is stated clearly within the assessment brief and supported by the self-assessment form. There may also be variation in how educators interpret and apply the relevant criteria. The Developmental model, meanwhile, depends on students receiving sufficient guidance and formative feedback to reflect meaningfully on their own practices. The language of the four pillars may also require further scaffolding for students who are unfamiliar with reflective writing or abstract academic terminology (Sadikoglu, 2025).

There are further limitations to the present paper. The Toolkit is introduced as a conceptual and design-based contribution, and no claims are made about its effectiveness. Its influence on student understanding, transparency, academic behaviour or learning has not been examined here. The self-assessment form also relies on students' accounts of their own AI use and should not be regarded as an independent method of verifying how work was produced or whether academic misconduct has occurred. Evidence from mandatory AI-use declarations suggests that fear of academic repercussions, ambiguous guidance and inconsistent enforcement may discourage accurate disclosure (Gonsalves, 2025). The form should therefore support reflection and transparency, but should not be considered as a complete evidential record. Similarly, the

practicality of the feedback resources may vary according to class size, marking workload and staff confidence. Where the four pillars contribute to formal marking through the Additive or Embedded models, consistency of interpretation, moderation and equitable application across educators also remain open questions.

The limitations discussed above provide the basis for future evaluation, which should examine not only whether the Toolkit can be implemented, but how it is understood and adapted in practice. Relevant questions include whether students interpret the four pillars consistently, whether the reflection prompts generate substantive rather than formulaic responses, and how educators distinguish between thoughtful reflection and simple disclosure. Evaluation should also examine how consistently the Additive and Embedded models are applied, whether the Developmental model supports more meaningful reflection over time and how students experience the different uses of the self-assessment form. Research should further consider the manageability of the feedback approaches, equitable access where AI use is required, how justified non-use is addressed in contexts where AI remains optional, and the adaptations required across levels and disciplines. A process of implementation, review and revision would allow these issues to inform the further development of the Toolkit, rather than assuming that it represents a finished solution (McKenney & Reeves, 2012). In this sense, the Toolkit provides a foundation for further pedagogical inquiry, while its educational value remains a matter for empirical examination.

## Conclusion

This paper has argued that academic integrity in AI-mediated assessment cannot be addressed solely by determining whether AI was used. The more important educational task is to make the purpose, quality and consequences of that use visible within the assessment process. Building on Visin's (2025) AIAI Framework, the AI Academic Integrity Toolkit translates Ethical Access, Critical Competence, Reflexive Transparency and Educational Value into three adaptable routes: Additive, Embedded and Developmental. These routes are supported by assessment brief language, student self-assessment and feedback resources designed to help educators connect expectations concerning AI use with the learning outcomes students are expected to demonstrate (Sadikoglu, 2025). The Toolkit does not treat AI use as a universal assessment requirement and recognises that, where it remains optional, students may make a reasoned decision not to use it. Its contribution is to provide a clear structure through which students can justify their choices, evaluate AI-generated outputs and reflect on whether technological assistance has supported genuine learning. Whether these aims are realised must be examined through implementation and empirical evaluation. Nevertheless, by locating integrity within the design of assessment rather than leaving it primarily to post-submission detection, the Toolkit offers a practical foundation for a more transparent, critical and educational response to generative AI in higher education.

## Acknowledgements

The author would like to thank colleagues at Anglia Ruskin University London for the professional discussions that informed the development of the AI Academic Integrity Toolkit.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Grammarly, an AI-assisted writing software, was used in proofreading and refining the language used in the manuscript. The usage included correcting grammatical and spelling errors and rephrasing statements for accuracy, clarity and readability. The author further declares that no other AI or AI-assisted technologies have been used in writing the manuscript. The Toolkit design, selection and interpretation of the literature, analysis and discussion presented in the manuscript are the author's own.

## References

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university: What the student does* (4th ed.). McGraw-Hill Education.

Bretag, T. (2013). Challenges in addressing plagiarism in education. *PLOS Medicine*, *10*(12), e1001574. https://doi.org/10.1371/journal.pmed.1001574

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Fawns, T. (2019). Postdigital education in design and practice. *Postdigital Science and Education*, *1*(1), 132–145. https://doi.org/10.1007/s42438-018-0021-8

Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, *1*, 3–31.

Gonsalves, C. (2025). Addressing student non-compliance in AI use declarations: Implications for academic integrity and assessment in higher education. *Assessment & Evaluation in Higher Education, 50*(4), 592–606. https://doi.org/10.1080/02602938.2024.2415654

International Center for Academic Integrity. (2021). *The fundamental values of academic integrity* (3rd ed.). https://academicintegrity.org/aws/ICAI/asset_manager/get_file/911282?ver=1

Kofinas, A. K., Tsay, C. H.-H., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology, 56*(6), 2522–2549. https://doi.org/10.1111/bjet.13585

McKenney, S., & Reeves, T. C. (2012). *Conducting educational design research.* Routledge.

O'Dea, X., & O'Dea, M. (2023). Is artificial intelligence really the next big thing in learning and teaching in higher education? A conceptual paper. *Journal of University Teaching and Learning Practice*, *20*(5), 1–17.

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice*, *21*(6), 49–66. https://doi.org/10.53761/q3azde36

Perkins, M., Roe, J., & Furze, L. (2025). Reimagining the artificial intelligence assessment scale: A refined framework for educational assessment. *Journal of University Teaching and Learning Practice*, *22*(7), 1–26. https://doi.org/10.53761/rrm4y757

Price, M., Rust, C., O'Donovan, B., Handley, K., & Bryant, R. (2012). *Assessment literacy: The foundation for improving student learning*. Oxford Centre for Staff and Learning Development.

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching, 6*(1), 342–363. https://doi.org/10.37074/jalt.2023.6.1.9

Sadikoglu, R. (2025). A *practical toolkit for ethical and reflective assessment design in the age of AI* [Preprint]. figshare. https://doi.org/10.6084/m9.figshare.30340876

Visin, G. (2025). *AIAI – AI-enabled academic integrity* [Preprint]. figshare. https://doi.org/10.6084/m9.figshare.28827620.v1

**Contact email:** rahme.sadikoglu@london.aru.ac.uk

# AI Integration in Pedagogical Practices: A Quantitative Analysis of the Competence of College Teachers

Joan P. Guinto, University of the Cordilleras, Philippines
Shayne Klarisse E. Dinamling, University of the Cordilleras, Philippines
Ibrahim F. Hanbal, University of the Cordilleras, Philippines
Don Harl C. Malabanan, University of the Cordilleras, Philippines
Dionisio R. Tandingan Jr., University of the Cordilleras, Philippines

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

In the last few years, Artificial Intelligence (AI) became a driving force for globalization and economic shifts. Particularly, its impact on the educational sector became a major consideration with the paradigm shift of many educational institutions around the world. The purpose of this study is to examine the extent of competence of teachers with the integration of AI tools in their pedagogical practices. Through quantitative descriptive design, a survey questionnaire was administered to 388 college teachers. It was found that the college teachers are competent in terms of the use of various AI tools. This result is supported by best practices such as the provision of AI-focused faculty development programs and establishment of an AI team in the institution.

*Keywords:* AI integration, AI tools, AI use, pedagogical practices, college teachers

iafor
The International Academic Forum
www.iafor.org

## Introduction

In the last few years, Artificial Intelligence (AI) became a driving force for globalization and economic shifts. Particularly, its impact on the educational sector became a major consideration with the paradigm shift of many educational institutions around the world. It became a disruptive innovation that changed various educational processes particularly with teaching-learning activities such as assessment and content development.

## Literature Review

Generally, AI is defined as a technology that enables machines to perform with human-like abilities (Sheikh et al., 2023). Due to its fast development, it has been applied widely in educational practices such as intelligent tutoring systems, adaptive learning systems, and personalized learning (Kamalov et al., 2023; Ouyang & Jiao, 2021). Specifically, many teachers began to reconsider its integration towards their own pedagogical practices such as in their teaching practices and classroom management (Huang et al., 2021).

Even before it became popular in 2023, AI has been integrated in different facets of education for the last 15 years. Zhai et al. (2021) analyzed a total of 100 research articles from 2010 and found that there are three main dimensions in which AI is used in education. They described these three dimensions as development (i.e.: classification, recommendation, and deep learning), extraction (i.e.: feedback, reasoning, and adaptive learning), and application (i.e.: role playing, immersive learning, and gamification). In addition, Holmes and Tuomi (2022) identified three foci on how AI is commonly used in various educational institutions namely: student-focused AI (i.e.: AI-assisted simulations, chatbots, and dialogue-based tutoring systems), teacher-focused AI (i.e.: plagiarism detection, smart curation of learning materials, and automatic summative assessment), and institution-focused AI (i.e.: admissions, course planning, and school security).

Despite its numerous potentials, AI integration in education still presents issues and challenges. The popularity of generative AI such as ChatGPT presented issues on ethical and responsible use of AI in education (Bibi et al., 2024). Issues in data privacy and transparency present questions on the collection and analysis of vast amounts of student data (Bibi et al., 2024; Chima et al., 2024).

Similarly, the Philippines also faces issues about data privacy concerns. However, unlike other countries, there are also other pressing problems that should be addressed such as digital divide, unequal access, and lack of faculty training and development (Dalan, 2023; Estrellado & Miranda, 2023). Hence, this study aims to address these research gaps through exploring the familiarity and competence of college teachers in integrating AI in their pedagogical practices.

This study sought to examine how college teachers of the University of the Cordilleras integrate Artificial Intelligence (AI) tools in their pedagogical practices. Specifically, this study sought to answer the following questions:

1. What is the extent of familiarity of college teachers with the use of AI tools?
2. What is the extent of competence of college teachers with the integration of AI tools in their pedagogical practices?

## Methodology

This study employed descriptive quantitative research design to examine the familiarity of college teachers with the use of AI tools and their extent of competence with the integration of AI tools in their pedagogical practices.

A two-part survey questionnaire entitled Artificial Intelligence Training Needs Assessment (AI-TNA) Tool was developed. The first part consists of 6 indicators focusing on the familiarity of the teachers with AI tools whereas the second part consists of 9 indicators focusing on the level of competence of teachers in AI integration. This survey questionnaire was distributed online through the use of Google form.

The study was conducted at the University of the Cordilleras, Philippines. The survey questionnaire was administered to 388 college teachers from 10 colleges, namely: College of Accountancy, College of Architecture and Fine Arts, College of Arts and Sciences, College of Business Administration, College of Criminal Justice Education, College of Engineering, College of Hospitality Tourism and Management, College of Information Technology and Computer Sciences, College of Nursing and Allied Health Services, and College of Teacher Education.

Before the data gathering collection, the researchers requested the Vice-President of Academics for the approval to float the AI-TNA Tool to the various colleges. During the data gathering collection, the researchers provided a letter of consent for the respondents for them to know the basic ethical standards applied in the study. After the data gathering collection, the collected data were tallied, analyzed, and interpreted.

Weighted mean was used to identify the extent of familiarity of college teachers with the use of AI tools and their extent of competence with the integration of AI tools in their pedagogical practices. Table 1 presents the extent of familiarity of college teachers with the use of AI tools whereas Table 2 presents the extent of competence of college teachers with the integration of AI tools in their pedagogical practices.

**Table 1**
*Statistical Range on the Extent of Familiarity of College Teachers With the Use of AI Tools*

| Statistical Range | Descriptor | Interpretation |
|---|---|---|
| 3.25–4.00 | Very Familiar | The college teachers have comprehensive understanding and knowledge about the use of AI tools. |
| 2.50–3.24 | Familiar | The college teachers have adequate understanding and knowledge about the use of AI tools. |
| 1.75–2.49 | Slightly Familiar | The college teachers have limited understanding and knowledge about the use of AI tools. |
| 1.00–1.74 | Not Familiar | The college teachers have little to no knowledge about the use of AI tools. |

**Table 2**
*Statistical Range on the Extent of Competence of College Teachers With the Integration of AI Tools in Their Pedagogical Practices*

| Statistical Range | Descriptor | Interpretation |
|---|---|---|
| 3.25–4.00 | Highly Competent | The college teachers demonstrate comprehensive knowledge, skills, and abilities in AI tools integration. |
| 2.50–3.24 | Competent | The college teachers demonstrate adequate knowledge, skills, and abilities in AI tools integration. |
| 1.75–2.49 | Slightly Competent | The college teachers demonstrate limited knowledge, skills, and abilities in AI tools integration. |
| 1.00–1.74 | Not Competent | The college teachers demonstrate little to no knowledge, skills, and abilities in AI tools integration. |

The researchers were given a notice to proceed by the ethics committee of the University of the Cordilleras. In their letter of consent, the researchers informed their respondents about data privacy and confidentiality. They assured the respondents that the data will solely be used for the purpose of the study and that the respondents will not receive remuneration for their participation.

## Results and Discussion

This part of the study presents the results, interpretation, and discussion of the data collected. Table 3 presents the analysis of data on the extent of familiarity of college teachers with the use of AI tools whereas Table 4 presents the extent of competence of college teachers with the integration of AI tools in their pedagogical practices.

**Table 3**
*Extent of Familiarity of College Teachers With the Use of AI Tools*

| AI Tools | Weighted Mean | Descriptive Equivalent |
|---|---|---|
| Generative Tools (i.e.: ChatGPT, Google Gemini, and Microsoft Copilot) | 3.46 | Very Familiar |
| Grammar Checking Tools (i.e.: Grammarly and QuillBot AI) | 2.97 | Familiar |
| AI for Data Analysis (i.e.: Excel Copilot, ChatGPT, and MathGPT) | 2.90 | Familiar |
| Creative Tools (i.e.: Canva Ai and Microsoft Designer) | 2.88 | Familiar |
| AI-Assisted Research Tools (i.e.: NotebookLM, Elicit, ChatPDF, and Research Rabbit) | 1.78 | Slightly Familiar |
| Learning Management Systems (LMS) with AI Features (i.e.: Magic School AI and Khanmigo) | 1.78 | Slightly Familiar |
| **General Weighted Mean** | **2.81** | **Familiar** |

Generally, the college teachers are familiar (2.81) with the use of AI tools. This means that the teachers have adequate understanding and knowledge about the use of AI tools. This result supports the study of Kurshumova (2024) in which the majority of Bulgarian teachers have higher levels of familiarity with AI technology. According to College Teacher 9, the university's Institute of Teaching and Learning conducted a series of AI capacity building workshops that helped her become more familiar with the AI tools mentioned in the survey questionnaires. In addition, College Teacher 20 specifically mentioned programs such as AI institutional trainings and programs such as Technological Consortium for Awareness, Readiness, and Advancement Knowledge in Innovation (TARAKI) led by the Department of Science and Technology (DOST) and the university's Office of the Vice-President for Research and Innovation.

Specifically, the college teachers are very familiar (3.46) with the use of generative tools (i.e.: ChatGPT, Google Gemini, and Microsoft Copilot). Most of the teachers use generative tools for the preparation of reports (College Teacher 1), creation of syllabus (College Teacher 9), and development of learning materials (College Teacher 12). Based on these statements, the use of generative tools has indeed increased throughout the years as they are perceived as valuable tools not only for teachers but also for students (Bibi, Yamin, Natividad, Rafique, Akhter, Fernandez, & Samad, 2024; Kaplan-Rakowski, Grotewold, Hartwick, & Papin, 2023).

On the other hand, the college teachers are only slightly familiar (1.78) with the use of LMS with AI features (i.e.: Magic School AI and Khanmigo). Despite the initiatives of the university, there are still some teachers who do not have training with the use of AI features in the LMS. Others like College Teacher 1 mentioned that some of the AI workshops in relation to the use of LMS are only introductory. In the Philippines, Heretape and Paglinawan (2025) enumerated challenges such as limited digital literacy, unstable internet connection, and sustainability of

student engagement as factors hindering the effective use of LMS and the integration of other AI and digital tools.

**Table 4**
*Extent of Competence of College Teachers With the Integration of AI Tools in Their Pedagogical Practices*

| AI Competence Indicators | Weighted Mean | Descriptive Equivalent |
|---|---|---|
| I use AI ethically and responsibly in teaching and learning. | 3.27 | Highly Competent |
| I use AI-powered learning tools for academic productivity. | 3.22 | Competent |
| I apply my understanding on AI's impact on education and the workplace. | 3.09 | Competent |
| I use AI tools to assist in content creation. | 3.07 | Competent |
| I understand what AI is and its general application. | 3.05 | Competent |
| I evaluate the content generated by AI tools in terms of accuracy, bias, and reliability. | 3.05 | Competent |
| I promote digital and AI literacy among students. | 2.97 | Competent |
| I facilitate personalized learning and student engagement using AI tools. | 2.89 | Competent |
| I use AI to assist in research, data analysis, or publication support. | 2.88 | Competent |
| **General Weighted Mean** | **3.02** | **Competent** |

Generally, the college teachers are competent (3.02) with the integration of AI tools in their pedagogical practices. This means that the college teachers demonstrate adequate knowledge, skills, and abilities in AI tools integration. In congruence, College 6 narrated that she integrates various AI tools in the courses that she handles. In his courses, College Teacher 18 exemplified his competence in AI integration as he uses AI tools in demonstrating how to use AI in creating and re-creating models and prototypes. Furthermore, College Teacher 16 emphasized that integrating AI to her pedagogical practices provides varied activities for different learning styles. Similar to these experiences, Ren and Wu (2025) argued that the competence of college teachers in AI technologies enables a more meaningful teaching-learning process such as AI-supported innovative pedagogy and AI-driven learning experience design and delivery.

Specifically, the college teachers are highly competent (3.27) with using AI ethically and responsibly in teaching and learning. Prioritizing ethical principles in AI integration promotes fairness, transparency, accountability, and inclusivity that can support the diverse needs and aspirations of students (Abbas, 2023; Chima, Onyebuchi, & Idowu, 2024). The ethical and responsible use of AI was also highlighted by College Teacher 2 as she discussed to the students

about the "careful" use of AI; however, despite this ethical awareness, she added that the creation of a standardized policy on AI use within the university is also needed. College Teacher 3 and College Teacher 9 also noted that a guideline on proper, ethical, and responsible use of AI is an important addition to AI integration.

On the other hand, the result shows that the lowest indicator is the college teachers' use of AI to assist in research, data analysis, or publication support (2.88). Despite their competence on the said indicator, many of the college teachers such as College Teachers 1, 2, 3, 6, 7, and 15 mentioned that they need more training with the use of AI in research. In addition, College Teachers 1 and 19 also cited paid subscriptions and other funded support for the use of AI are needed to strengthen the competence of the college teachers in the integration of AI in research.

## Conclusion

The introduction of AI is one of the most transformative integrations of technology in the field of education in the 21st century. With this change, teachers are expected to be the forerunners of AI as they themselves try to navigate this innovation. In this study, college teachers demonstrated both familiarity and competence in the use and integration of AI tools in their pedagogical practices. Initiatives such as capacity building through the university's AI team provided the needed assistance for the college teachers in exploring various AI tools. As a result, the college teachers were able to use and integrate AI tools in various teaching and learning contexts. However, university support such as the creation of AI guidelines and subscription to various AI tools are still needed by the college teachers to further their knowledge and abilities in using and integrating AI tools in their teaching practices.

## Acknowledgement

The researchers are grateful to God for the wisdom and strength to finish this study. They also extend their gratitude to the University of the Cordilleras for the support provided throughout the course of writing this study. They would also like to thank their families for the love, support, and prayers that they have received. *Iyaman!*

## References

Abbas, T. (2023). Ethical implications of AI in modern education: Balancing innovation and responsibility. *Social Sciences Spectrum, 2*(1), 51–57.

Bhana, A. (2024). Unlocking the power of convergent parallel designs and triangulation for enhanced management and leadership research: A comprehensive theoretical exploration. *Asian Journal of Management, Entrepreneurship, and Social Science, 4*(04), 1770–1793.

Bibi, A., Yamin, S., Natividad, L., Rafique, T., Akhter, N., Fernandez, S., & Samad, A. (2024). Navigating the ethical landscape: AI integration in education. *Educational Administration: Theory and Practice, 30*(6), 1579–1585. https://doi.org/10.53555/kuey.v30i6.5546

Chima, A., Onyebuchi, N., & Idowu, S. (2024). Integrating AI in education: Opportunities, challenges, and ethical considerations. *Magna Scientia Advanced Research and Review, 10*(2), 6–13. https://doi.org/10.30574/msarr.2024.10.2.0039

Estrellado, C., & Miranda, J. (2023). Artificial intelligence in the Philippine educational context: Circumspection and future inquiries. *International Journal of Scientific and Research Publication, 13*(5), 16–22. http://dx.doi.org/10.29322/IJSRP.13.04.2023.p13704

Heretape, E., & Paglinawan, J. (2025). Tenured instructors' views on the use of learning management system (LMS) in online classes. *International Journal of Scientific and Management Research, 8*(10), 343–351. http://doi.org/10.37502/IJSMR.2025.81027

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, Research, Development, and Policy, 57*(4), 542–570. https://onlinelibrary.wiley.com/doi/abs/10.1111/ejed.12533

Huang, J., Saleh, S., & Liu, Y. (2021). A review of artificial intelligence in education. *Academic Journal of Interdisciplinary Studies, 10*(3), 206–217. https://pdfs.semanticscholar.org/4590/d37ca3f650e9f72613189003a8c49eddb75b.pdf

Kamalov, F., Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability, 15*(16), 12451. https://doi.org/10.3390/su151612451

Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., & Papin, K. (2023). Generative AI and teachers' perspectives on its implementation in Education. *Journal of Interactive Learning Research, 34*(2). https://doi.org/10.70725/815246mfssgp

Kurshumova, D. (2024). A snapshot of Bulgarian school teachers' familiarity with, use of, and opinions on artificial intelligence at the threshold of its incorporation into the educational process. *Discover Education, 3*, 138.

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence, 2*, 100020. https://doi.org/10.1016/j.caeai.2021.100020

Piedad, E. (2024). Strengthening artificial intelligence-on-edge education in the Philippines: A teacher-centric curriculum development. *2024 IEEE 13th International Conference on Engineering Education (ICEED),* 1–6. https://doi.org/10.1109/ICEED62316.2024.10923809

Ren, X., & Wu, M. (2025). Examining teaching competencies and challenges while integrating artificial intelligence in higher education. *Tech Trends, 69,* 519–538. https://doi.org/10.1007/s11528-025-01055-3

Sheikh, H., Prins, C., & Schrijvers, E. (2023). Artificial intelligence: definition and background. In *Mission AI. Research for Policy* (pp. 15–41). Springer. https://doi.org/10.1007/978-3-031-21448-6_2

Zhai, X., Chu, X., Chai, C., Jong, M., Istenic, A., Spector, M., Liu, J., Yuan, J., & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity, 1*, 8812542. https://doi.org/10.1155/2021/8812542

**Contact emails:** jopucya@uc-bcf.edu.ph
sedinamling@uc-bcf.edu.ph
ifhanbal@uc-bcf.edu.ph
dcmalabana@uc-bcf.edu.ph
drtandingan@uc-bcf.edu.ph

# Time Series Forecasting Using Context Rich Restaurant Sales Data in Data Science Education

Boyang Zhang, University of Turku, Finland

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Restaurant sales forecasting in data science education is significant for decision making, human resource management, and inventory control. This paper proposes a decision-centric framework which includes context layer, data layer, model layer, and decision layers for the design and implementation of context-aware forecasting model. The Kaggle's Restaurant Sales Report 2024–2025 dataset is used in this research, it provides comprehensive daily sales records across multiple restaurant categories, enriched with contextual variables such as weather, promotions, special events, and pricing information. The methodology of this paper implements a structured analytical pipeline consisting of data cleaning, feature construction, time-series decomposition (trend and seasonality), and exploratory regression analysis using contextual variables such as weather, promotions, events, and prices. These steps are mapped to the proposed context, data, model, and decision layers to demonstrate how analytical outputs inform forecasting-related decisions. The findings of this research suggest that the proposed framework is suitable for forecasting the restaurant sales to practical decision making, by explicitly various layers, the framework helps to understand how data preprocessing, feature construction and modeling choices affect downstream outcomes, the prediction of the sales for the future. The framework supports the implementation of forecasting models that are transparent, reproducible and adaptable to various restaurant settings.

*Keywords:* data science education, time series forecasting, ensemble learning

iafor
The International Academic Forum
www.iafor.org

## Introduction

Nowadays, food waste in restaurants has become a global challenge, and it carries significant economic and environmental consequences. In order to understand the scale of the issue, recent studies estimate that the food service industry is responsible for approximately 26.1% of global food waste (Vardopoulos et al., 2025). Food waste brings negative influence on the daily running of restaurants; therefore, sales forecasting is vital important to inventory management, human resource management and sustainability management. Accurate sales forecasting is a challenging task, as reliable forecasting enables restaurants to optimize food purchases, staff schedules, pricing and even table allocations (Aulya, 2024; Boomija et al., 2018; Tanizaki et al., 2019). Difficulty arises because restaurant environments are dynamic, and with complex periodicities (A. Zhao & Jayadi, 2021; Tanizaki et al., 2020), the sales can fluctuate sharply by the time of year, month, week, holidays, and hours of a day (Güler et al., 2024; Žunić et al., 2020). At the same time, factors like weather, promotions, events as the various of context can also exert profound influence of sales in restaurants.

In order to make accurate forecasting, this paper presents a multi-layer decision centric process for sales prediction. It introduces a four-layer architecture which includes: context layer; data layer; model layer and decision layer. The four-layer architecture is designed to decompose data inputs and outputs, by structure data preprocessing, feature construction, modeling choices, and decision-making.

A key strength of this four-layer architecture is its universal application across most of the restaurant settings. Because this architecture can adjust its variables based on customized data, like local holidays, special events, building occupancies, and so on. It can be fitted with different business scales and dining concepts. With feature engineering of the historical dataset, the architecture can dynamically select the best model for prediction, and use the optimal weighting strategy for customized data. And with time gone by, the architecture will keep on updating and learning from historical data, and adjust the model and weight to make the most practical predictions. For example, the architecture provides its model layer with basic statistics analysis plus the latest analysis from non-linear variables. By combining predictive accuracy with computational efficiency, this four-layer architecture provides reliable, data-driven guidance for strategic business decision-making.

## Theoretical Background and Proposed Framework

The concept of Information System (IS) theory is an interdisciplinary field connected with the collection, processing, storage and support for organization decision making (Davis, 1974). And Information System theory provides a diagram for understanding how organizations are able to capture complex contexts information and interpret it into organizational strategies (Hananiya et al., 2024). Especially in the food industry, organizational strategies are connected to the precise synchronization of production (Aulya, 2024), inventory management (Edmond Yeboah et al., 2022), and human resource management (Aich et al., 2026). The restaurant industry is characterized by the simultaneous production and consumption of perishable services and food materials. Inappropriate planning might leads to strategic failures, including massive food waste, suboptimal service quality, and extra labor costs (Koivunen et al., 2020; Vardopoulos et al., 2025). In order to avoid above risks, restaurant management requires an Information System that builds balance between external contexts and internal organizational strategies.

Restaurant sales data constitute a complex time series characterized by temporal dependencies, seasonality, and multiple contextual influences (Aulya, 2024). Variables such as weather conditions, localized special events, pricing strategies, and promotional campaigns introduce are difficult as the nonlinear disruptions to standard time series forecasting (Boomija et al., 2018; Panda & Mohanty, 2023). Traditional time series models (e.g., ARIMA) are difficult to capture these high variance anomalies. Moreover, machine learning algorithms (e.g., XGBoost, LightGBM), while adept at mapping nonlinearity, are prone to overfitting noisy historical data, with limited predictive accuracy (Aydın et al., 2025). Accurate forecasting remains challenging because of the complex temporal dynamics of restaurant sales data.

In order make accurate restaurant sales forecasting, this paper proposes a decision-centric framework which includes context layer, data layer, model layer, and decision layers. It is a four-layer context aware forecasting model for mapping the data flow of Information System from the real-world context, through data acquisition and processing, to decision support.

**Figure 1**
*Four-Layer Decision Centric Framework*

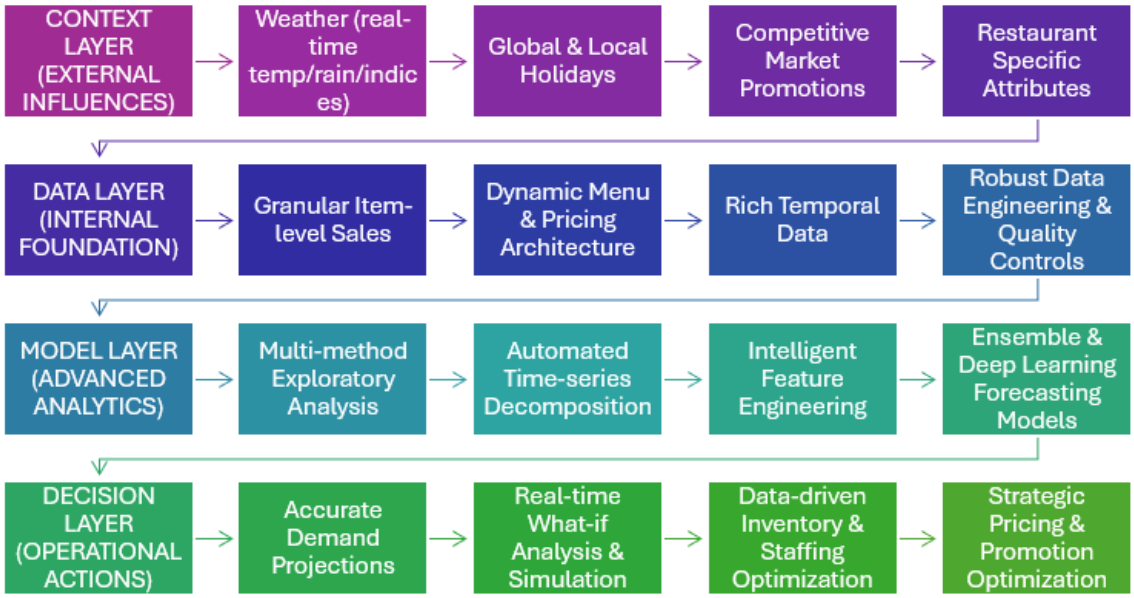



Figure 1 is the overall details of the four-layer decision centric frame. The first layer is the context layer. In IS theory, an open system must continuously interact with its external environment to remain viable and accurate (Davis, 1974; Hananiya et al., 2024). The context layer serves as the system's perceptual boundary, and this layer defines the multidimensional environment that the forecasting model needs to accommodate, encompassing variables such as real-time weather conditions which can include temperature and rain indices; global and local holidays; market promotions; and restaurant specific attributes which can include purchasing prices of product, human resource and so on. The data layer shows that system requires to collect possible external influences which might have an influence on sales.

The second layer is the data layer, it includes the storage of data and how data is cleaned. Because it is time series forecast, it needs to parse data into telemetry format. The granular item-level sales data emphasizes the use of transactional log data. This log data flows into dynamic menu and pricing architecture, such as the shifting of raw materials costs, menu

updates, and promotion pricing. The relevant data sources are integrated as the rich temporal data, ensuring every transactional and financial metric is anchored with timestamp as the telemetry data. The entire second layer is about data engineer to clean the data, and make the data clean enough and get ready for next steps.

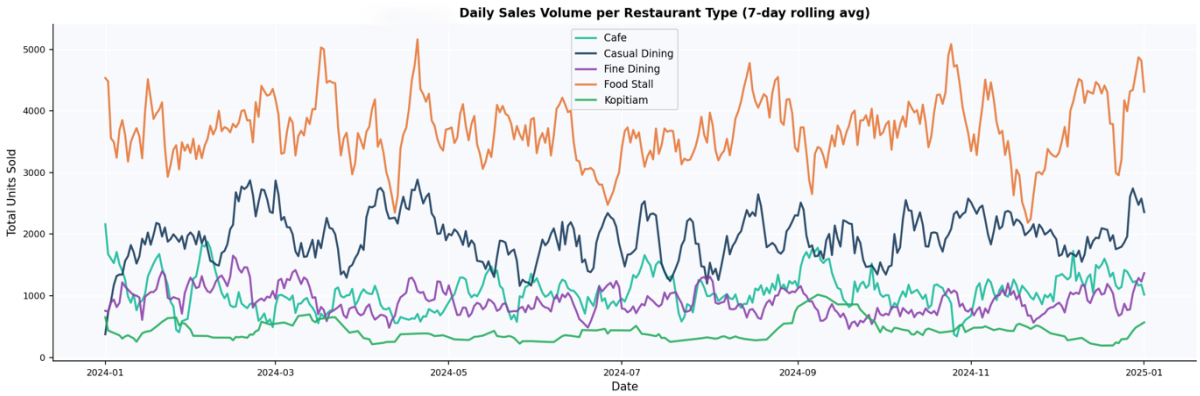
As the core of the system, the third model layer, can be regarded as the computational engine of the system. From the perspective of Information System, the computational model functions as the critical uncertainty reduction mechanism within the forecasting architecture (Taipalus et al., 2020). Multiple models are tested and evaluated based on the dataset. And a hybrid ensemble methodology could be used to make predictions based on accuracy from different models. The model layer is engineered to capture deterministic baseline patterns.

The last layer is the decision layer, it uses analytical outputs from model layer, and translates the predictions into corporate operation strategies. By directly informing critical organizational proposals, such as workforce scheduling, inventory optimization, and supply chain procurement. This layer interprets the prediction into suggestions on practical organizational decision making.

## Kaggle Restaurant Dataset

In this chapter, the dataset is introduced, and decision-centric framework are discussed. The study uses the open public Kaggle Restaurant Sales Report 2024–2025 dataset, which provides an item level view of daily sales across fifty diverse restaurants over a full year, from January 2024 to January 2025. The establishments span five categories including cafes, casual dining, fine dining, food stalls, and kopitiams.

**Figure 2**
*Daily Item Level Sales Volume per Restaurant Type*

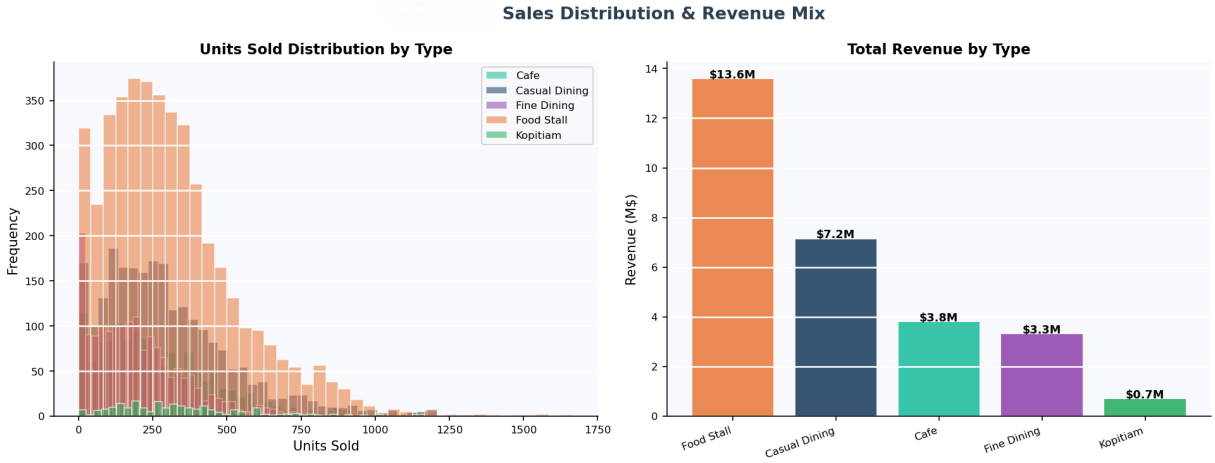



In order to get an overall view of the dataset, Figure 2 offers a longitudinal analysis of unit sales across five categories from January 2024 to January 2025. By utilizing a 7-day rolling average, it shows weekly and seasonal trends across the period of time. Food Stalls (represented by the orange line) are the undisputed leaders in sales volume. It fluctuated by sharp peaks in the springtime between March to April, and August as the beginning of school, alongside a noticeable mid-summer gap around July, and another peak before December. Casual Dining (dark blue line) maintains a solid and distinct second place, with total units sold generally fluctuating between 1,500 and 3,000. While its volume is substantially lower than that of Food Stalls, Casual Dining exhibits similar cyclical demand

patterns throughout the year. The remaining three categories, Cafes, Fine Dining, and Kopitiams, operate at considerably lower sales volumes, also with fluctuations.

**Figure 3**

*Sales Distribution and Revenue Mix*



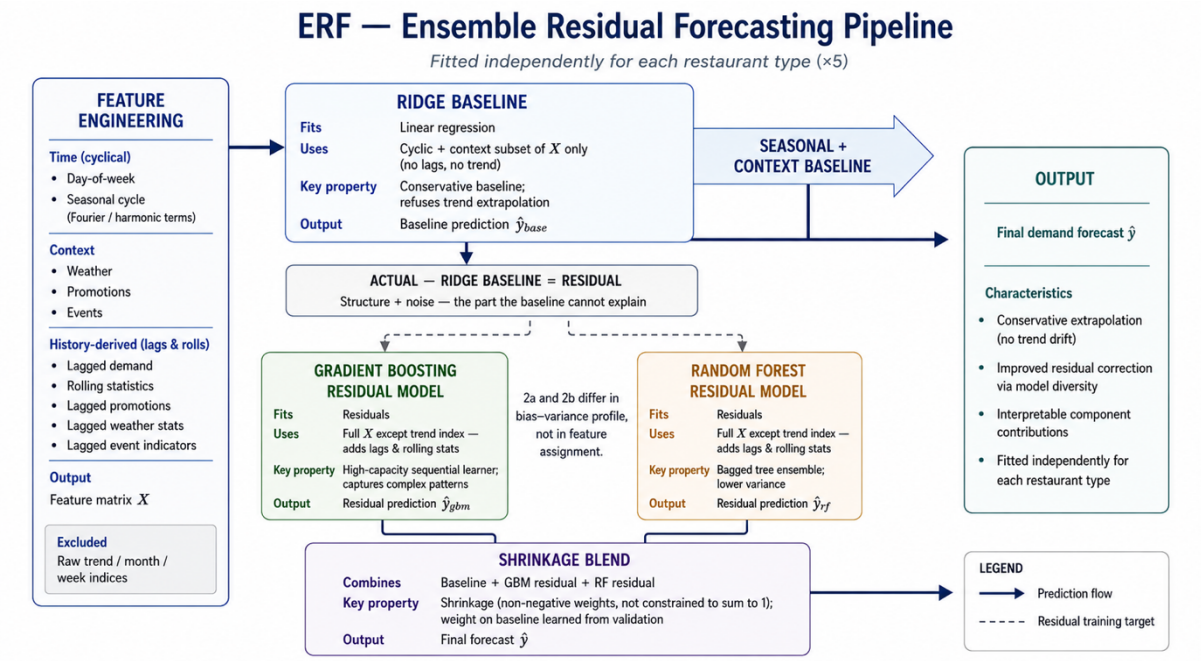

As to the sales distribution and revenue mix, we can see from Figure 3. The units sold distribution by restaurant type graphic, it shows a histogram detail of the frequency of daily units sold. The "Frequency" axis captures the exact number of restaurant days that achieved specific sales volumes across the dataset, 50 specific restaurants using daily temporal variables. Food Stalls dominate the distribution with the highest occurrence rates and the widest variance, recording days with well over 500 units sold and extending into a long tail past 1,250 units. Casual Dining forms the secondary volume tier, peaking in the 100 to 400 unit range. The remaining categories, which include Cafes, Fine Dining, and Kopitiam packed at the far left of the distribution, confirming that these establishments record lower daily transaction volumes, rarely exceeding 500 units on any given restaurant-day.

In the right graphic of Figure3, it is the analysis of "Total Revenue by Type" which shows a bar chart quantifying the cumulative financial output for each category in millions of dollars. Aligning with its sheer dominance in sales volume, the Food Stall category leads the revenue mix at $13.6M, almost doubling the earnings of the category of Casual Dining, which generated $7.2M. Cafes and Fine Dining follow with closely matched total revenues of $3.8M and $3.3M, respectively, highlighting how higher price points in fine dining can offset lower unit volumes to achieve revenue parity with higher volume cafes. Moreover, Kopitiam accounts for the smallest share of the financial mix, contributing $0.7M.

## The Hybrid Model: Ensemble Residual Forecasting (ERF)

The hybrid model, Ensemble Residual Forecasting (ERF) is designed based on the four decision centric layers, it fosters the development of models that are transparent, reproducible, and adaptable to diverse restaurant settings. The concept of Ensemble Residual Forecasting represents a sequential pipeline, from the statistical baseline of linear factors to the dynamical machine learning model for nonlinear variance. The residual acts as the error or variance for optimization and updating the weighting strategies of different models.

**Figure 4**
*ERF Ensemble Residual Forecasting Pipeline*



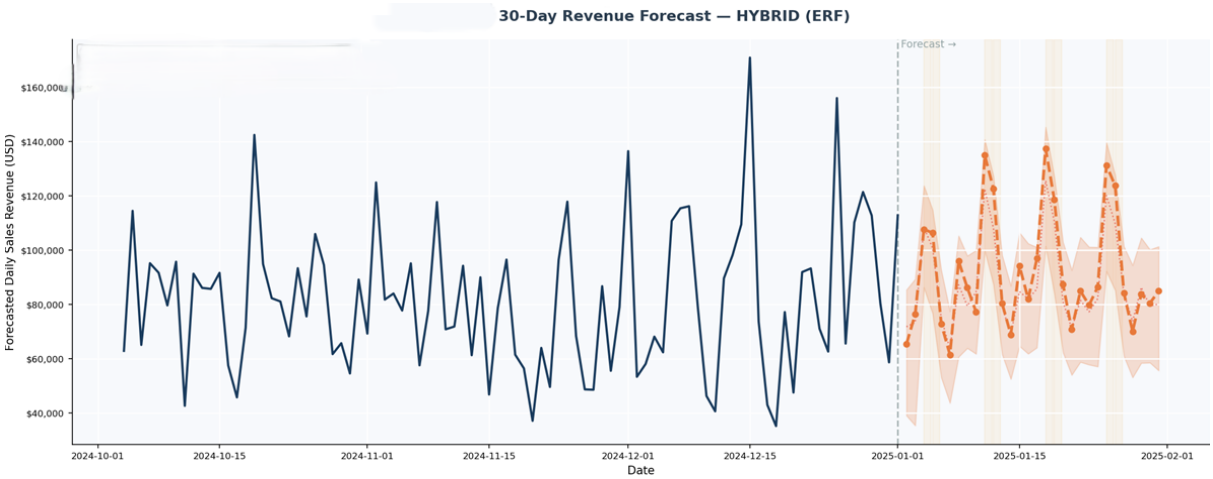

In the context layer, it begins with feature engineering of the Kaggle Restaurant Sales Report 2024–2025 dataset, which carries thirteen fields: date, restaurant_id, restaurant_type, menu_item_name, meal_type, key_ingredients_tags, typical_ingredient_cost, observed_market_price, actual_selling_price, quantity_sold, has_promotion, special_event, and weather_condition. There are multiple ways of classification of above variables, for example, three factors are engineered into daily context features: weather_condition (Sunny / Cloudy / Rainy) is encoded into (+1 / 0 / −1) and averaged per day; for has_promotion and special_event, boolean values are used. All fields in the dataset are categorized into variables for further analysis.

In the data layer, it serves as the foundational layer for building the comprehensive matrix X. It structures raw inputs into three distinct dimensions: cyclical time (day-of-week, Fourier terms) to encode stable seasonality, contextual variables (weather, promotions, events) to capture exogenous shocks, and history-derived telemetry (lagged demand, rolling statistics) to establish autoregressive momentum. Data layer is not fixed for different cases, in this research, the data layer is specifically designed for the Kaggle Restaurant Sales Dataset. It can be different in other cases, such as adding additional resources, like traffic flow variable, parking variable. In this case, for example, from the menu and pricing fields (menu_item_name, meal_type, key_ingredients_tags, typical_ingredient_cost, observed_market_price, and actual_selling_price), they are parsing into numeric financial signals. Categorical item details are mathematically encoded to represent menu composition, while continuous pricing data is transformed into elasticity metrics.

In the model layer, there are three stages in the process. First, a Ridge regression sets a stable, drift-free baseline using the cyclical and contextual features. Second, gradient boosting and random forest algorithms act as correctors; they are trained on the baseline's residuals, the errors, to reduce bias and variance. Finally, a shrinkage blend combines these outputs: the final forecast is built by taking the baseline prediction and adding the specific corrections from both the gradient boosting and random forest models, each scaled by its own learned

weight. The weight strategies are dynamic and are updating based on the new input data as with the time gone by.

**Figure 5**
*30-Day Revenue Forecast, the Hybrid Model*




The decision layer acts as the system's final uncertainty reduction mechanism, grounding the mathematical forecasts in Information System. It interprets the hybrid outputs from the stable baseline and the specific residual corrections into actionable business intelligence. The forecast is delivered for next 30 days. By providing clear, interpretable signals with prediction, this layer empowers restaurant managers to optimize daily operations like inventory and staffing, successfully bridging the gap between predictive analytics and strategic decision-making. In Figure 5, we provide the forecast for next 30 days with a range prediction, for example one day's sale can between 85.000 to 115.000 dollar. Individual different types of restaurants can have different sales predictions based on previous sales data. Whatmore, restaurant managers can make monthly purchase of products plan. In terms of practical business deployment and model evaluation design, this study adopts a dynamic rolling strategy. To ensure that the system does not deviate from the actual sales, the framework incorporates a next day update mechanism: whenever new day's actual sales data is generated, the system automatically updates the lag features and recalculates the residual correction weights to forecast future sales. Therefore, this will offer guidance for restaurant managers who can adjust the weekly plan and monthly plan accordingly.

## Discussion and Conclusion

In this chapter, a decision centric framework for restaurant sales forecasting is presented by utilizing a hybrid model, the Ensemble Residual Forecasting (ERF) model. There are four layers in the framework including context, data, model and decision layers, which decompose the data pipeline into structured process based on the Information System theory. It aims to make accurate forecasts on restaurant sales when encountering multiple linear and non-linear features.

In the evaluation, the hybrid model achieved optimal performance across key metrics. In Table1, the hybrid model's Mean Absolute Error (MAE) dropped to 126.6, meaning the daily sales forecast deviates by only about 127 units (meals/items). Combined with a Weighted Average Percentage Error (WAPE) of 11.4% and the highest explained variance ($R^2 = 0.476$), this variance is manageable for buffer stock control, proving the model's exceptional practical

viability compared to other models. The evaluation demonstrates the effectiveness of the hybrid ERF model's design in applying the four-layer framework. The framework provides the data pipeline, and the ERF model provides practical solutions and interprets mathematical variables to business insights.

**Table 1**
*Evaluation Results of the Models*

| Model | MAE | RMSE | $R^2$ | WAPE (%) |
|---|---|---|---|---|
| Naive Baseline | 302.6 | 359.2 | 1.209 | 27.4 |
| GBM (standalone) | 143.1 | 178.8 | 0.343 | 12.3 |
| Ridge (statistical) | 301.6 | 355.6 | 1.286 | 28.0 |
| Stacking Ensemble | 144.7 | 174.8 | 0.247 | 12.6 |
| HYBRID (ERF) | 126.6 | 162.0 | 0.476 | 11.4 |

While the four layer frameworks with ERF model demonstrate good performance, this study has several limitations that should be addressed. In the Kaggle Restaurant Sales Report dataset, it includes diverse operational contexts in five restaurant types, there could be more restaurant types, like holiday restaurants or institutional restaurants. Other types of restaurant data could also be tested. In order to filter out extreme daily noise, the ERF model's target variable was designed as a 7-day rolling average of units sold. While this is effective for weekly inventory procurement and staff scheduling, it masks sudden, single day sales spikes. If restaurant managers are looking for hourly or exact daily peak predictions, it would update the single day sales to hourly sales or in minutes sales, to catch up the extreme peaks for different angles of predictions. Moreover, the current pipeline forecasts total daily demand aggregated by restaurant type. It does not investigate the demand for individual menu items or specific raw ingredients, which would require handling even higher levels of data sparsity and noise.

In conclusion, this research has provided a general framework which can be applied across any restaurant settings, from local food stalls to fine dining chains. It delivers an accurate forecasting architecture that empowers restaurant managers with the early planning and decision-making capabilities based on different contexts, and it can help restaurant managers to optimize daily inventory, streamline labor scheduling, and reduce food waste. The four-layer framework is able to dynamically adjust its correction weights based on users' data. Future research should expand the ERF architecture to forecast demand for individual menu items, which would allow the system to generate direct, itemized procurement lists for kitchens. Additionally, the current ERF architecture relies on structured contextual features like weather and promotions, future iterations could integrate more non-linear variables, such as customer online feedback or social media sentiments and emotions, to better capture sudden changes in sales.

## References

A. Zhao, M., & Jayadi, R. (2021). *Forecasting Daily Visitors and Menu Demands in an Indonesian Chain Restaurant using Support Vector Regression Machine,* 1–6. https://doi.org/10.1109/AIMS52415.2021.9466036

Aich, M., Kushwah, R., Bisui, R., & Tiwari, S. (2026). Strategic HR Interventions for Workforce Excellence in the Food Production Sector: A Study of Best Practices and Outcomes. *Advances in Consumer Research. Association for Consumer Research (U.S.)*, *3*, 350–359.

Aulya, N. (2024). Production Capacity Planning Strategy to Deal with Seasonal Demand in the Food and Beverage Industry. *Maroon Journal De Management*, *1*, 132–142. https://doi.org/10.37899/mjdm.v1i3.99

Aydın, B., Balcioglu, Y., & Sezen, B. (2025, March). *Food demand forecasting with machine learning methods: An application using decision tree*.

Boomija, G., Anandaraj, A., Nandhini, S., & Lavanya, S. (2018). Restaurant Visitor Time Series Forecasting Using Autoregressive Integrated Moving Average. *Journal of Computational and Theoretical Nanoscience*, *15*, 1590–1593. https://doi.org/10.1166/jctn.2018.7345

Davis, G. B. (1974). *Management Information Systems: Conceptual Foundations, Structure, and Development*. McGraw-Hill.

Edmond Yeboah, N., Opoku, R., & Kaku, G. (2022). Inventory strategies and performance of food and beverage processing industries. *International Journal of Logistics Systems and Management*, *41*, 120. https://doi.org/10.1504/IJLSM.2022.120985

Güler, A., Musa, A., Tarım, M., Saraç, O., & Gokturk, M. (2024). Forecasting Restaurant Sales with the Sensitivity of Weather Conditions and Special Days Using Facebook Prophet. *Journal of Data Applications*, 15–30. https://doi.org/10.26650/JODA.1450459

Hananiya, P., Badau, K., Olajire, F., & Yidawi, M. (2024). *Information systems in organizations*.

Koivunen, L., Rauti, S., Laato, S., Naskali, J., Nissilä, P., Ojansivu, P., Mäkilä, T., & Norrdal, M. (2020). Increasing Customer Awareness on Food Waste at University Cafeteria with a Sensor-Based Intelligent Self-Serve Lunch Line. *Proceedings of the IEEE International Conference on Smart Computing (SMARTCOMP)*. https://doi.org/10.1109/TMC49519-2030.9138571

Panda, S., & Mohanty, S. (2023). Time Series Forecasting and Modeling of Food Demand Supply Chain Based on Regressors Analysis. *IEEE Access*, *11*, 42679–42700. https://doi.org/10.1109/ACCESS.2023.3266275

Taipalus, T., Seppänen, V., & Pirhonen, M. (2020). Uncertainty in information system development: Causes, effects, and coping mechanisms. *Journal of Systems and Software*, *168*, 110655. https://doi.org/https://doi.org/10.1016/j.jss.2020.110655

Tanizaki, T., Hoshino, T., Shimmura, T., & Takenaka, T. (2019). Demand forecasting in restaurants using machine learning and statistical analysis. *Procedia CIRP*, *79*, 679–683. https://doi.org/10.1016/j.procir.2019.02.042

Tanizaki, T., Hoshino, T., Shimmura, T., & Takenaka, T. (2020). Restaurants store management based on demand forecasting. *Procedia CIRP*, *88*, 580–583. https://doi.org/https://doi.org/10.1016/j.procir.2020.05.101

Vardopoulos, I., Abeliotis, K., & Lasaridi, K. (2025). A Systematic Informetric Analysis and Literature Review of Food Waste Quantification Studies in the Food Service Industry. *Sustainability*, *17*(1). https://doi.org/10.3390/su17010103

Žunić, E., Korjenić, K., Hodžić, K., & Donko, D. (2020). Application of Facebook's Prophet Algorithm for Successful Sales Forecasting Based on Real-World Data. *International Journal of Computer Science & Information Technology (IJCSIT)*, *12*(2), 23–36. https://doi.org/10.5121/ijcsit.2020.12203

# Critical Reflection and Inclusive Education: A Semester-Long Implementation of Brookfield's Four Lenses in Teacher Education

Natassa Raikou, University of Thessaly, Greece
Labrina Gioti, Aristotle University of Thessaloniki, Greece

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Inclusive education requires teachers who are able not only to apply inclusive pedagogical strategies but also to critically examine the beliefs and assumptions that shape their professional practice. However, teacher education has devoted considerably more attention to developing pedagogical knowledge than to systematically fostering critical reflection on the implicit assumptions underlying inclusive teaching. In response to this gap, the present pilot study reports the implementation of a semester-long reflective intervention based on Brookfield's Four Lenses of Critical Reflection within an experiential, workshop-based undergraduate course for pre-service special education teachers. Forty undergraduate students from the Department of Special Education at the University of Thessaly participated in the intervention. Throughout the semester, students engaged in a sequence of individual and collaborative reflective activities organised around Brookfield's four complementary perspectives: autobiographical experience, learners' perspectives, colleagues' perspectives, and theoretical literature. Qualitative data derived from reflective writings, lesson plans, collaborative activities and group discussions were analysed using thematic analysis. The findings indicate that participants progressively questioned previously taken-for-granted assumptions about teaching and inclusion, developed broader understandings of inclusive pedagogy, and increasingly connected critical reflection with their future professional practice. Reflective activities appeared to support participants in recognising inconsistencies between declared beliefs and pedagogical thinking while encouraging more inclusive and critically informed interpretations of educational practice. Although exploratory in nature, this pilot implementation suggests that Brookfield's Four Lenses provide a promising pedagogical framework for embedding systematic critical reflection within initial teacher education. The study contributes practical insights into the design of reflective learning environments that may support the preparation of future teachers for inclusive educational practice.

*Keywords:* critical reflection, Brookfield's four lenses, inclusive education, teacher education, pre-service teachers, thematic analysis

iafor
The International Academic Forum
www.iafor.org

## Introduction

Inclusive education has become one of the central priorities of contemporary educational policy and teacher education worldwide. International organisations, including UNESCO (2020) and the European Agency for Special Needs and Inclusive Education, consistently emphasise that developing inclusive schools requires not only structural reforms but also teachers who are able to respond critically, ethically, and flexibly to increasingly diverse classrooms. Consequently, teacher education is expected to prepare future educators who are capable of making pedagogical decisions that promote participation, equity, and meaningful learning for all students.

Although substantial progress has been made in identifying the knowledge, skills, and attitudes associated with inclusive teaching, research increasingly suggests that these competencies alone may not be sufficient. Numerous studies indicate that teachers' pedagogical decisions are strongly influenced by deeply held beliefs, implicit assumptions, and previous educational experiences, many of which remain unexamined during initial teacher education (Florian & Black-Hawkins, 2011; Walton & Rusznyak, 2020). As a result, future teachers may express positive attitudes towards inclusion while simultaneously reproducing instructional practices that unintentionally restrict participation or reinforce exclusion.

This challenge has directed increasing scholarly attention towards critical reflection as a central component of teacher education. Rather than being understood simply as a reflective skill or a professional competence, critical reflection enables teachers to identify, question, and reconstruct the assumptions that shape their pedagogical reasoning and professional judgement (Brookfield, 2017; Mezirow, 2000). From this perspective, reflective practice becomes not merely an individual cognitive activity but a transformative pedagogical process through which teachers critically reinterpret experience, integrate multiple perspectives, and develop more inclusive ways of understanding teaching and learning.

Among the various approaches to critical reflection, Stephen Brookfield's Four Lenses provides one of the most comprehensive frameworks for examining teaching from multiple perspectives. By integrating autobiographical experience, learners' perspectives, colleagues' feedback, and theoretical knowledge, the model encourages educators to make implicit assumptions visible and open them to critical examination (Brookfield, 2017). Although Brookfield's framework has been widely recognised within adult and higher education, relatively limited empirical research has explored its systematic implementation as a semester-long pedagogical approach in pre-service teacher education, particularly within the context of inclusive education.

The present study is grounded in the assumption that critical reflection is not simply an educational competence but the pedagogical mechanism through which inclusive professional thinking develops. Rather than examining reflection as an isolated educational activity, the study investigates the implementation of Brookfield's Four Lenses as the organising principle of an experiential university course designed to support critically reflective learning. Through a qualitative analysis of students' reflective writings, lesson plans, collaborative activities, and group discussions, the study explores how systematic engagement with multiple reflective perspectives contributes to the reconstruction of pre-service teachers' pedagogical thinking about inclusive education.

More specifically, the study addresses the following research questions:

RQ1. How does systematic engagement with Brookfield's Four Lenses support pre-service teachers in recognising and questioning implicit assumptions about inclusive education?

RQ2. How do participants' understandings of inclusive education evolve throughout the semester-long reflective intervention?

RQ3. In what ways do participants connect critical reflection with their emerging professional practice?

## Theoretical Framework

### Inclusive Education and Teachers' Beliefs

Inclusive education is increasingly understood as a process of creating educational environments in which all learners can participate meaningfully, experience a sense of belonging, and develop according to their individual strengths and needs (Ainscow, 2020; UNESCO, 2020). Contemporary perspectives move beyond the traditional emphasis on disability or placement and instead conceptualise inclusion as a continuous process of identifying and removing barriers to participation and learning (Ainscow, 2020). Consequently, learner diversity is regarded not as an obstacle but as an educational resource that enriches teaching and learning (Florian, 2014; Florian & Black-Hawkins, 2011).

Within this perspective, inclusive pedagogy extends beyond the implementation of differentiated instructional techniques. It requires teachers to exercise informed professional judgement in designing learning environments that maximise participation for all learners while avoiding practices that unintentionally reinforce exclusion (Florian & Black-Hawkins, 2011; Messiou, 2017). As Florian and Black-Hawkins (2011) argue, inclusive teaching is characterised by extending learning opportunities to everybody rather than creating separate opportunities for particular groups of students.

However, research consistently demonstrates that the successful implementation of inclusive pedagogy depends not only on teachers' professional knowledge but also on the beliefs through which they interpret learner diversity and make pedagogical decisions (Dignath et al., 2022; Fives & Buehl, 2012; Pajares, 1992). Teachers' beliefs influence how they understand students' abilities, evaluate classroom situations, and respond to diversity in everyday educational practice. Consequently, two teachers with similar professional preparation may enact substantially different forms of inclusive teaching because they interpret educational situations through different belief systems.

An important challenge, however, is that many of these beliefs operate implicitly. Teachers frequently remain unaware of the assumptions developed through their own educational experiences and professional socialisation, despite the considerable influence these assumptions exert on classroom practice (Brookfield, 2016, 2017). This observation is consistent with Argyris and Schön's (1974) distinction between *espoused theories* and *theories-in-use*, according to which educators' declared professional beliefs do not always correspond to the assumptions that actually guide their pedagogical decisions. Consequently, supporting inclusive education requires more than promoting positive attitudes or providing pedagogical knowledge; it also requires creating opportunities for future teachers to identify, question, and reconstruct the assumptions underlying their professional judgement.

## Critical Reflection as the Pedagogical Mechanism of Inclusive Teacher Education

If teachers' beliefs frequently remain implicit, an important question emerges: how can future teachers become aware of the assumptions that shape their professional judgement? Contemporary research increasingly identifies critical reflection as one of the most powerful pedagogical processes for addressing this challenge. Rather than focusing solely on the acquisition of knowledge or technical teaching skills, critical reflection encourages educators to examine the beliefs, values, experiences, and social influences that underpin their pedagogical decisions (Brookfield, 2017; Fook & Gardner, 2007; Gioti, 2019).

Within teacher education, critical reflection extends beyond descriptive accounts of experience. It involves questioning what is often taken for granted, recognising the influence of personal biographies and institutional cultures, and examining how dominant assumptions shape understandings of learners, teaching, and participation (Brookfield, 2016, 2017; Gioti, 2019; Gioti & Toka, 2025). Such reflection enables future teachers to move from asking "What *happened?"* towards asking *"Why do I interpret educational situations in this particular way?"* and *"Which assumptions guide my pedagogical decisions?"*

This understanding aligns closely with transformative learning theory, according to which meaningful professional learning occurs when individuals critically reassess previously unquestioned frames of reference and reconstruct them through dialogue, experience, and critical examination (Mezirow, 2000). Transformation therefore represents neither the acquisition of new information nor the simple modification of attitudes. Rather, it involves a gradual reconstruction of the perspectives through which professional situations are interpreted. Previous research has similarly demonstrated that sustained opportunities for critical reflection support greater self-awareness, empathy, and professional growth in teacher education (Raikou, 2016, 2019; Ryan, 2015).

However, critical reflection rarely develops spontaneously. Deeply rooted assumptions are seldom challenged through isolated reflective assignments or occasional classroom discussions. Instead, reflective learning requires structured pedagogical environments that provide repeated opportunities for dialogue, feedback, perspective-taking, and the continuous revisiting of assumptions over time (Brookfield, 2017; Gioti, 2024; Ryan, 2015). Consequently, the quality of teacher education depends not only on what future teachers learn but also on how reflective learning is pedagogically organised.

Among the available approaches to critically reflective teaching, Stephen Brookfield's Four Lenses offers one of the most comprehensive frameworks for systematically examining professional assumptions. Brookfield proposes that educators critically analyse their practice through four complementary perspectives: their own autobiographical experiences, learners' perspectives, colleagues' feedback, and theoretical literature (Brookfield, 2017). Each lens contributes differently to the reflective process, while their combined use increases the likelihood of identifying assumptions that would otherwise remain invisible.

A central contribution of Brookfield's framework lies in the interaction among these four perspectives. Personal experience alone is inevitably subjective; learners provide valuable but partial perspectives; colleagues challenge professional routines through dialogue; and theory offers conceptual distance that enables educators to reinterpret experience within broader educational and social contexts. Reflection therefore becomes increasingly rigorous when these

perspectives are brought into dialogue rather than considered independently (Brookfield, 2017; Gioti, 2024).

This multi-perspective approach aligns closely with contemporary understandings of inclusive pedagogy. Inclusive education requires teachers to examine assumptions concerning ability, participation, difference, and educational success while remaining responsive to learners' diverse experiences and identities (Florian, 2014; Walton & Rusznyak, 2020). Brookfield's framework provides the pedagogical process through which such assumptions may become visible, questioned, and gradually reconstructed. In this sense, critical reflection functions not simply as a professional competence but as the pedagogical mechanism connecting inclusive educational values with professional judgement and classroom decision-making.

Accordingly, the present study examines the systematic implementation of Brookfield's Four Lenses throughout an entire academic semester, investigating how sustained engagement with multiple reflective perspectives contributes to the development of critically reflective and inclusive pedagogical thinking among pre-service teachers.

## Methodology

### Research Design

The study adopted a qualitative interpretive design to explore how systematic engagement with critical reflection contributes to the development of inclusive pedagogical thinking among pre-service teachers. Rather than measuring changes in predetermined variables, the study sought to understand how participants gradually reconstructed their assumptions, interpretations, and professional reasoning throughout a semester-long educational intervention. This interpretive orientation was considered particularly appropriate because the research focused on the developmental processes through which critically reflective learning unfolds in teacher education.

The intervention was designed as a semester-long reflective learning environment structured around Stephen Brookfield's Four Lenses. Instead of treating critical reflection as an isolated educational activity or a final reflective assignment, the Four Lenses constituted the organising pedagogical framework of the entire course. The intention was to provide participants with repeated opportunities to examine educational experiences through multiple perspectives, revisit previously held assumptions and progressively reconstruct their understanding of inclusive education.

### Participants and Educational Context

The study was conducted during the winter semester of the academic year 2025–2026 at the Department of Special Education of the University of Thessaly. Forty undergraduate pre-service teachers voluntarily participated in the intervention as part of a compulsory laboratory-based course within their teacher education programme. The course combined theoretical input, experiential activities, collaborative learning, reflective writing, lesson planning, and structured dialogue, allowing students to engage continuously with Brookfield's autobiographical, learners', colleagues', and theoretical lenses throughout the semester.

Rather than following a linear sequence of independent teaching sessions, the intervention was intentionally organised as a recursive reflective process in which participants repeatedly

revisited previous experiences and interpretations from different perspectives. This iterative design aimed to support increasingly sophisticated forms of critical reflection and inclusive pedagogical reasoning.

## Educational Intervention

The intervention was developed around four complementary reflective perspectives derived from Brookfield's Four Lenses. Initially, participants critically engaged with theoretical literature concerning inclusive education, transformative learning, and critical reflection, providing the conceptual foundations for subsequent reflective work. They were then invited to revisit significant autobiographical educational experiences, encouraging awareness of the influence of previous schooling on present pedagogical beliefs. Experiential learning activities and simulations subsequently enabled students to explore educational situations from learners' perspectives, while collaborative dialogue and peer feedback encouraged critical examination of alternative interpretations. Finally, participants designed inclusive lesson plans, implemented teaching activities, and reflected individually and collectively on their pedagogical decisions. Importantly, these learning experiences should not be understood as separate instructional stages. Throughout the semester students continuously revisited previous assumptions, integrating new perspectives into an ongoing process of reflective reconstruction.

## Data Collection

Multiple qualitative sources were collected in order to capture participants' reflective development from complementary perspectives. The dataset included:

- reflective journals,
- written reflective assignments,
- inclusive lesson plans,
- collaborative classroom activities,
- group discussions.

Collecting different forms of evidence enabled comparison between students' reflective narratives, pedagogical planning, collaborative interactions, and evolving interpretations throughout the intervention.

## Data Analysis

The data were analysed using Braun and Clarke's (2006) thematic analysis. Rather than focusing on the frequency of particular responses, the analysis sought to identify recurring patterns that reflected changes in participants' ways of thinking about inclusive education.

Particular attention was paid to evidence indicating:

- increasing awareness of implicit assumptions;
- movement from descriptive towards critical reflection;
- reconstruction of pedagogical reasoning;
- development of empathy and perspective-taking;
- emerging understandings of inclusive pedagogy.

The iterative analysis involved repeated reading of the data, initial coding, development of candidate themes, refinement of thematic categories, and continuous comparison across

different qualitative sources. Because evidence was collected throughout the semester and across different learning activities, the analysis aimed not only to identify the content of students' reflections but also to trace patterns of reflective development across the semester.

The methodological contribution of the study lies not only in implementing Brookfield's Four Lenses but in embedding them within the pedagogical architecture of an entire university course. This design enabled the investigation of reflective development as an ongoing educational process rather than as the outcome of isolated reflective activities.

## Findings

The thematic analysis identified three interrelated findings that describe participants' reflective development throughout the semester-long intervention. Rather than representing isolated outcomes, these findings suggest a gradual shift in how pre-service teachers understood inclusive education and their future professional role. Across the different stages of the intervention, participants moved beyond describing educational experiences towards questioning previously taken-for-granted assumptions, reconstructing their understanding of inclusion, and reflecting more critically on their own pedagogical responsibilities. The following sections present each finding together with representative excerpts from participants' reflective writings.

### Questioning Previously Taken-for-Granted Assumptions

One of the clearest findings concerned participants' growing awareness of previously unexamined assumptions about teaching and inclusion. At the beginning of the intervention, many students described inclusive education primarily in normative terms, emphasising equality, acceptance and support for learners with disabilities. However, engagement with autobiographical reflection, critical incidents and theoretical perspectives encouraged them to reconsider how their own educational experiences had shaped their beliefs about learners, participation and classroom practice.

Rather than simply recalling past experiences, participants increasingly analysed them critically, recognising how seemingly ordinary classroom routines could unintentionally reinforce exclusion or limit participation. Reflection therefore became less descriptive and increasingly interpretive, enabling students to question assumptions they had previously considered self-evident.

As one participant explained: "I realised that many teaching practices I had always considered normal had actually made me feel anxious as a learner. This made me reflect on how easily I could reproduce similar practices without even realising it."[1]

These reflections suggest that critical reflection functioned not simply as a process of recalling experience but as a means of making implicit assumptions visible and open to critical examination.

[1] All participant excerpts were translated from Greek into English by the authors.

## Reconstructing Understandings of Inclusive Education

A second important finding concerned the gradual reconstruction of participants' understanding of inclusion. Initially, inclusion was often associated with adapting instruction for specific groups of learners or providing additional support to students with disabilities. Throughout the intervention, however, students increasingly described inclusion as a broader pedagogical orientation characterised by participation, belonging, learner voice and equitable learning opportunities for all students.

Experiential activities, perspective-taking tasks and the learners' lens appeared particularly influential in this process. Rather than focusing exclusively on instructional adaptations, participants increasingly emphasised the emotional, relational and social dimensions of inclusive teaching.

One participant reflected: "The activity helped me understand that learning difficulties are not only academic. They also affect students' emotions, confidence and participation in the classroom."

Similarly, another participant concluded: "Inclusion is not about making different activities for a few students. It is about creating a classroom where everyone feels respected, safe and able to participate."

These reflections indicate a gradual movement from technical interpretations of inclusion towards a more holistic understanding centred on participation, relationships and educational equity.

## Reflection as a Foundation for Future Professional Practice

The third finding concerns the way participants increasingly connected reflection with their emerging professional identity. By the end of the intervention, reflective writing moved beyond evaluating individual activities and focused more explicitly on participants' future role as teachers. Students frequently described critical reflection as an essential component of professional learning, emphasising the importance of continuously questioning their own pedagogical decisions, listening to learners' perspectives and engaging in dialogue with colleagues.

Collaborative reflection was consistently described as particularly valuable because it exposed participants to alternative interpretations of educational situations and challenged the certainty of their own viewpoints.

As one participant noted: "Through discussions with my classmates I realised that some practices I considered supportive could actually exclude students from active participation."

Another participant concluded: "This course changed the way I think about my future role as a teacher. I now understand that being an inclusive teacher means constantly reflecting on my own practice and being willing to reconsider my assumptions."

Overall, these findings suggest that participants increasingly perceived critical reflection not as an isolated academic exercise but as an ongoing professional responsibility closely connected with participants' future professional practice.

Taken together, these preliminary findings suggest that sustained engagement with multiple reflective perspectives may support pre-service teachers in moving from descriptive accounts of teaching towards more critically informed understandings of inclusive pedagogy. While the present paper reports the principal findings of this pilot implementation, further analysis of the complete dataset is expected to provide a more comprehensive understanding of the developmental processes underlying participants' reflective learning.

## Discussion

The present pilot study explored the implementation of a structured reflective intervention based on Brookfield's (2017) Four Lenses within an undergraduate teacher education course. Overall, the findings suggest that the intervention created meaningful opportunities for pre-service teachers to engage in progressively deeper forms of critical reflection while simultaneously broadening their understanding of inclusive education.

The first finding indicates that participants increasingly questioned assumptions that had previously remained implicit in their thinking about teaching and learning. Rather than merely describing educational experiences, students began to examine how their own beliefs, prior experiences and taken-for-granted understandings influenced their interpretations of classroom situations. This finding is consistent with Brookfield's (2017) argument that critical reflection begins with identifying assumptions that shape professional practice. It also resonates with Mezirow's (2000) view that examining previously unquestioned assumptions constitutes an essential condition for perspective transformation.

A second important outcome concerns participants' evolving understanding of inclusive education. Throughout the intervention, inclusion gradually shifted from being viewed primarily as instructional support for particular groups of learners towards a broader understanding centred on participation, belonging, learner voice and equitable learning opportunities. This development aligns with contemporary approaches to inclusive pedagogy, which emphasise creating learning environments that support the participation of all learners rather than focusing exclusively on individual differences (Florian, 2014; Florian & Black-Hawkins, 2011; Mansouri et al., 2024). The findings therefore suggest that reflective pedagogies may help future teachers move beyond procedural interpretations of inclusion towards more relational and democratic understandings of classroom practice. This interpretation is also consistent with recent evidence showing that teachers' understandings of inclusion significantly influence the quality of inclusive classroom practice (Woodcock & Anderson, 2025).

The third finding highlights the educational value of engaging with multiple perspectives throughout the reflective process. Reflection was strengthened not only through individual self-examination but also through dialogue with peers, engagement with educational theory and consideration of learners' perspectives. These findings support Brookfield's (2017) proposition that critical reflection is enhanced when teachers deliberately examine practice through multiple lenses rather than relying solely on personal experience. Participants frequently described peer dialogue and learner feedback as opportunities to reconsider their own interpretations, suggesting that collaborative reflection may represent an important component of teacher education.

Although the present study was designed as a pilot implementation, the findings indicate that structured reflective interventions can contribute to the development of reflective dispositions

that are closely connected with inclusive teaching, complementing recent evidence highlighting the importance of socio-emotional competencies for inclusive professional learning (Stavropoulou et al., 2026). Rather than presenting reflection as an isolated academic exercise, the intervention encouraged participants to view reflective inquiry as an ongoing aspect of professional learning and future teaching practice. This observation is consistent with previous research emphasising the importance of systematic reflection for supporting professional judgement, continuous professional learning, and inclusive pedagogical decision-making (Brookfield, 2017; Raikou, 2019; Ryan, 2015).

Several limitations should also be acknowledged. The study was conducted with a single cohort of pre-service teachers within one university context, and the findings are based exclusively on qualitative reflective artefacts produced during the intervention. Consequently, the results should be interpreted as exploratory rather than generalisable. Future research could examine similar interventions across different teacher education contexts and combine qualitative and quantitative approaches in order to investigate changes in reflective competence and inclusive professional learning over time. In addition, further analysis of the complete corpus of reflective data is expected to provide a more comprehensive understanding of how different reflective perspectives interact throughout teacher education and how they contribute to the development of inclusive professional learning.

## Conclusions and Implications

The present pilot study provides preliminary evidence that a structured reflective intervention based on Brookfield's Four Lenses can support critical reflection and promote more inclusive understandings of teaching among pre-service teachers. Across the intervention, participants demonstrated increasing willingness to question previously unexamined assumptions, consider multiple perspectives and reflect more critically on their future professional practice. These findings suggest that reflective activities grounded in authentic educational experiences can enrich teacher education by creating opportunities for meaningful dialogue, perspective-taking and professional self-examination.

The study also highlights the pedagogical value of combining autobiographical reflection, critical incidents, collaborative dialogue, learner feedback and theoretical perspectives within a coherent learning sequence. Rather than treating reflection as an isolated instructional technique, the intervention positioned reflective inquiry as a continuous process closely connected with inclusive pedagogy and professional learning. This finding supports current calls for teacher education programmes to embed structured opportunities for sustained reflection throughout both coursework and practicum experiences. More broadly, the findings underline the importance of designing teacher education not only as a process of acquiring pedagogical knowledge, but also as a process of critically examining assumptions, engaging with multiple perspectives and gradually developing reflective dispositions that support inclusive educational practice.

Given the exploratory nature of this pilot implementation, the findings should be interpreted with appropriate caution. Nevertheless, they indicate that Brookfield's Four Lenses constitute a promising framework for designing reflective learning environments that encourage pre-service teachers to critically examine educational practice from multiple perspectives. Future research should further investigate the long-term impact of such interventions across different teacher education contexts and examine more closely the developmental processes through which reflective practice contributes to inclusive professional learning.

## Acknowledgements

The research is conducted in the operating framework of the University of Thessaly Innovation, Technology Transfer Unit and Entrepreneurship Center "One Planet Thessaly," under the "University of Thessaly Grants for Scientific Publication Support" action and is funded by the Special Account of Research Grants of the University of Thessaly.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that Grammarly, an AI-assisted writing software, was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The authors further declare that, apart from Grammarly, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Ainscow, M. (2020). Inclusion and equity in education: Making sense of global challenges. *Prospects, 49*(3–4), 123–134. https://doi.org/10.1007/s11125-020-09506-w

Argyris, C., & Schön, D. A. (1974). *Theory in Practice: Increasing Professional Effectiveness*. Jossey-Bass.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brookfield, S. (2016). So exactly what is critical about critical reflection? In J. Fook, V. Collington, F. Ross, G. Ruch, & L. West (Eds.), *Researching Critical Reflection and Research: Multidisciplinary Perspectives* (pp. 23–34). Routledge.

Brookfield, S. D. (2017). *Becoming a critically reflective teacher* (2nd ed.). Jossey-Bass.

Dignath, C., Rimm-Kaufman, S., van Ewijk, R., & Kunter, M. (2022). Teachers' beliefs about inclusive education and insights on what contributes to those beliefs: A meta-analytical study. *Educational Psychology Review, 34*, 2609–2660. https://doi.org/10.1007/s10648-022-09695-0

Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook* (Vol. 2, pp. 471–499). American Psychological Association.

Florian, L. (2014). What counts as evidence of inclusive education? *European Journal of Special Needs Education, 29*(3), 286–294. https://doi.org/10.1080/08856257.2014.933551

Florian, L., & Black-Hawkins, K. (2011). Exploring inclusive pedagogy. *British Educational Research Journal, 37*(5), 813–828. https://doi.org/10.1080/01411926.2010.501096

Fook, J., & Gardner, F. (2007). *Practising critical reflection: A resource handbook*. Open University Press.

Gioti, L. (2019). Theories of Action and Theories-in-use of Adult Educators. Motivation, Obstacles and Individual Differences in Teachers' Training. *International Journal of Learning and Teaching, 5*(3), 233–238. https://doi.org/10.18178/ijlt.5.3.233-238

Gioti, L. (2024). *To ergo tou Stephen Brookfield: Theoria kai praktiki tis kritika stochastikis mathisis* [The work of Stephen Brookfield: Theory and practice of critically reflective learning]. Epistimoniki Enosi Ekpaidefsis Enilikon Publications.

Gioti, L., & Toka, K. (2025). From critical reflection to transformative praxis. Intersecting Freire, Mezirow & Brookfield. *Adult Education: Critical Issues, 5*(2), 51–65. https://doi.org/10.12681/haea.43677

Mansouri, M. C., Miller, A. L., Kurth, J. A., Ruhter, L., Wilt, C. L., Gross Toews, S., & Morningstar, M. E. (2024). Teachers' insights on cultivating inclusive education for students with complex support needs. *Teachers College Record, 126*(9). https://doi.org/10.1177/01614681241302637

Messiou, K. (2017). Research in the field of inclusive education: Time for a rethink? *International Journal of Inclusive Education, 21*(2), 146–159. https://doi.org/10.1080/13603116.2016.1223184

Mezirow, J. (2000). Learning to think like an adult. In J. Mezirow (Ed.), *Learning as transformation: Critical perspectives on a theory in progress* (pp. 3–33). Jossey-Bass.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307–332. https://doi.org/10.3102/00346543062003307

Raikou, N. (2016). Development of critical thinking through aesthetic experience: the case of students of an educational department. *Journal of Transformative Education, 14*(1), 53–70. https://journals.sagepub.com/doi/abs/10.1177/1541344615606535

Raikou, N. (2019). Teacher Education at the forefront: Long-term study through the prism of University Pedagogy and Transformative Learning theory. *European Journal of Education Studies, 6*(3), 88–102. https://doi.org/10.5281/zenodo.3243469

Ryan, M. E. (Ed.). (2015). *Teaching reflective learning in higher education: A systematic approach using pedagogic patterns*. Springer. https://doi.org/10.1007/978-3-319-09271-3

Stavropoulou, G., Beazidou, E., Raikou, N., Fragoulis, G., Siatras, A., & Sotiropoulos, D. (2026). Linking Inclusion and Self-Efficacy: The Mediating Role of Social-Emotional Learning in Pre-Service Teacher Education. *Journal of Research in Special Educational Needs, 26*, e70106. https://doi.org/10.1111/1471-3802.70106

UNESCO. (2020). *Global education monitoring report 2020: Inclusion and education – All means all*. UNESCO.

Walton, E., & Rusznyak, L. (2020). Cumulative knowledge-building for inclusive education in initial teacher education. *European Journal of Teacher Education, 43*(1), 18–37. https://doi.org/10.1080/02619768.2019.1686480

Woodcock, S., & Anderson, J. (2025). Conceptions to classrooms: The influence of teacher knowledge on inclusive classroom practice. *International Journal of Educational Research Open, 8*, 100412. https://doi.org/10.1016/j.ijedro.2024.100412

# A Multi-model Deliberative Generative AI Framework: Enhancing Critical Thinking and Content Veracity in Higher Education

Shyh-Jian Tan, WuFeng University, Taiwan
Y. C. K. Chen, WuFeng University, Taiwan

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Generative Artificial Intelligence (GAI), powered by Large Language Models (LLMs), has rapidly entered higher education, offered instant content generation but raising concerns about student overreliance and cognitive slacking. While GAI outputs often appear credible, they remain vulnerable to bias and hallucinations, creating risks of misinformation. Current teaching practices focus mainly on tool operation, lacking systematic approaches to guide students in verifying authenticity or engaging in critical evaluation. This study proposes the Multi-model Deliberative Generative AI Framework (MDGAIF), which employs cross-model verification and structured guidance to transform students from passive recipients into active evaluators and ethical stewards of AI. A quasi-experimental design was conducted with students in an "Introduction to Machine Learning" course, comparing an experimental group (MDGAIF-integrated teaching) with a control group (traditional tool-based teaching). Mixed-methods analysis revealed that MDGAIF significantly improved students' habits of verifying information and enhanced their critical thinking and problem-solving skills. Findings demonstrate that MDGAIF not only achieves its intended goals but also provides a valuable reference for universities seeking to design integrated AI teaching strategies and curricula.

*Keywords:* generative artificial intelligence, critical thinking, higher education, MDGAIF, content veracity

iafor
The International Academic Forum
www.iafor.org

## Introduction

The landscape of higher education has undergone a profound transformation with the advent of Generative Artificial Intelligence (GAI) technologies powered by Large Language Models (LLMs). Since the public release of ChatGPT in late 2022, followed by rapid advancements in competing platforms such as Google's Gemini, Anthropic's Claude, and Meta's Llama series, students have gained unprecedented access to AI systems capable of generating sophisticated text, code, images, and even video content with remarkable fluency and apparent expertise. This technological revolution has fundamentally altered the dynamics of academic work, enabling students to produce complex assignments, solve technical problems, and generate research content at speeds previously unimaginable.

### Research Context and Motivation

However, the convenience of Generative Artificial Intelligence (GAI) technologies have precipitated a critical pedagogical crisis that we term the "Cognitive Slacking" phenomenon. Students increasingly exhibit a pattern of over-reliance on AI-generated outputs, submitting content without rigorous verification, critical evaluation, or genuine intellectual engagement with the material. This behavioral pattern represents more than mere academic dishonesty; it constitutes a fundamental abdication of cognitive responsibility that threatens the core educational mission of developing independent critical thinking and problem-solving capabilities. The ease with which students can obtain plausible-sounding answers has created what Mollick and Mollick (2023) describe as a "delegation of thinking," where the cognitive labor of analysis, synthesis, and evaluation is outsourced entirely to algorithmic systems.

The risks associated with this uncritical adoption of GAI outputs extend beyond pedagogical concerns to encompass serious issues of content veracity and academic integrity. Despite their impressive linguistic capabilities, LLMs are fundamentally probabilistic systems that generate text based on statistical patterns in training data rather than genuine understanding or factual knowledge. This architectural limitation manifests in several critical failure modes: (1) *hallucinations*, where models confidently generate factually incorrect information that appears logically coherent; (2) *outdated information*, as models are trained on historical data and lack awareness of recent developments; (3) *biases*, reflecting problematic patterns in training data; and (4) *logical inconsistencies*, where reasoning chains contain subtle but significant errors (Zhang et al., 2025).

The emergence of so-called "reasoning models" in 2025—exemplified by OpenAI's o1 and o3 series (OpenAI, 2024) and Anthropic's Claude 3.7 Sonnet (Anthropic, 2025)—has paradoxically intensified these concerns. While these advanced systems demonstrate enhanced Chain-of-Thought (CoT) capabilities and can engage in more sophisticated multi-step reasoning, research indicates they remain vulnerable to producing "highly plausible hallucinations" (Zhang et al., 2025). The very sophistication of their outputs makes errors more difficult to detect, as they are embedded within otherwise coherent and logically structured responses. This creates what we term the "hallucination paradox": as models become more capable, their errors become more dangerous precisely because they are more convincing.

Furthermore, the recent shift toward "Agentic Workflows"—where AI systems can autonomously execute complex multi-step tasks with minimal human intervention—has amplified the cognitive slacking problem. When students deploy autonomous agents capable of independently researching, coding, writing, and debugging, the degree of automation often

triggers "deep cognitive delegation." Students become mere initiators and acceptors of AI-generated work, abandoning the essential quality audit processes that should accompany any intellectual production.

## Research Objectives

This study addresses the implementation gap by proposing and empirically validating the Multi-Model Deliberative Generative AI Framework (MDGAIF), a comprehensive pedagogical intervention designed to operationalize UNESCO's 2025 principles within the context of technical higher education. The framework transforms the student's role from that of a "prompt sender" who passively receives AI outputs to that of an "auditor and quality controller" who actively deliberates on, verifies, and takes responsibility for AI-assisted work.

The research pursues three primary objectives.

### *Objective 1: Framework Development*

To design and articulate a systematic, replicable pedagogical framework that embeds deliberative verification procedures directly into AI-assisted learning workflows, specifically tailored for technical education contexts where content veracity is paramount.

### *Objective 2: Empirical Validation*

To rigorously assess the framework's effectiveness through a quasi-experimental study measuring its impact on two critical dimensions: (a) critical thinking capabilities, as measured by standardized instruments adapted from Facione (1990) and Ng et al. (2021), and (b) complex problem-solving behaviors, as assessed through systematic analysis of student-AI interaction patterns based on OECD (2014) and Wing (2006) frameworks.

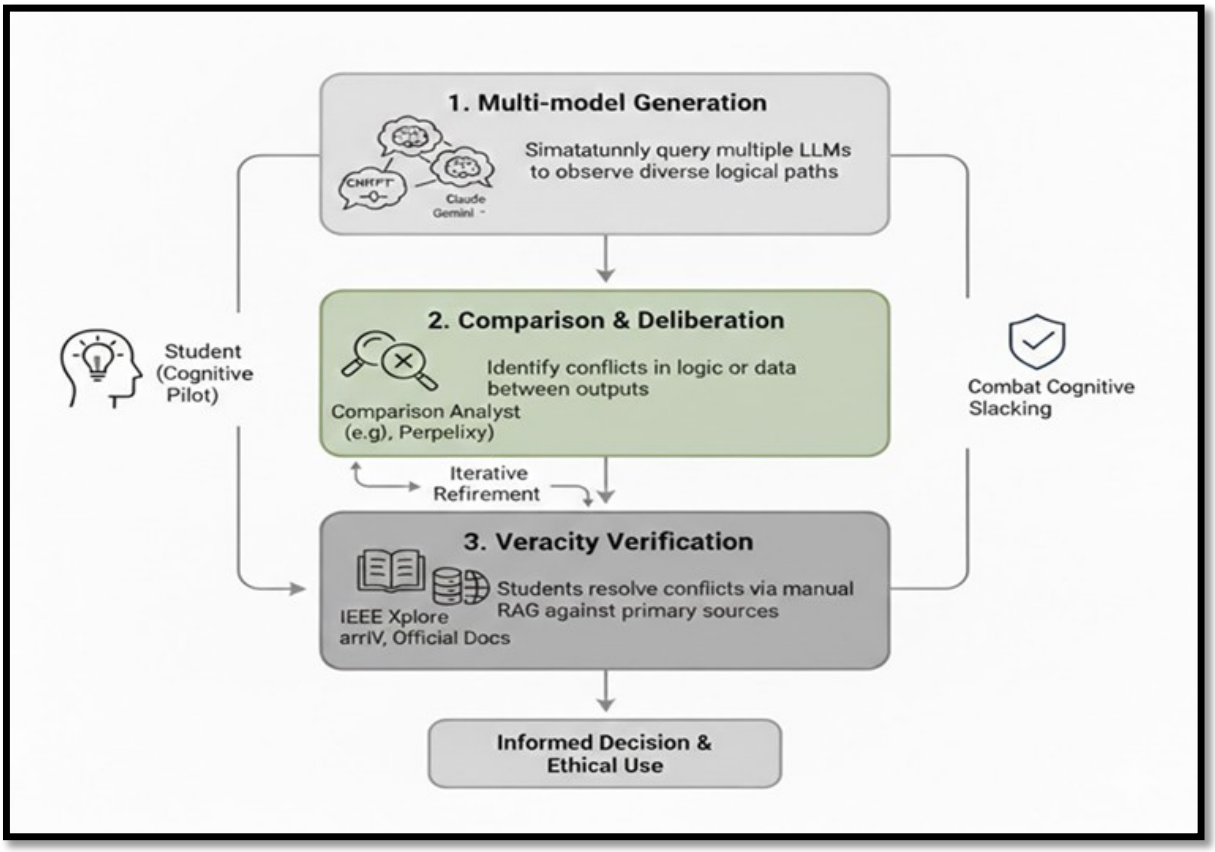
### *Objective 3: Pedagogical Translation*

To provide concrete implementation guidance, including specific instructional protocols, assessment rubrics, and operational procedures that enable other educators to adopt and adapt the framework within their own institutional contexts.

## MDGAIF Architecture and Process

The Multi-Model Deliberative Generative AI Framework (MDGAIF) transforms students from passive consumers into active deliberative evaluators. The MDGAIF architecture and processes are shown as Figure 1 which consists of three interconnected modules.

**Figure 1**
*Framework Architecture and Processes*




**Module 1: Multi-model Generation**

Students formulate a structured prompt and submit it to three distinct LLMs: ChatGPT, Claude, and Gemini. This induces "cognitive diversity," exposing students to multiple approaches and preventing premature acceptance of a single AI answer.

**Module 2: Comparison and Deliberative Conflict Identification**

Students use a fourth model, such as Perplexity, as a "Comparison Analyst." This model identifies consensus, conflicts in logic, and specific technical differences. This meta-level critical thinking teaches students to evaluate the AI's evaluative process.

**Module 3: Veracity Verification**

This represents the core deliberative component. Students resolve conflicts through "Manual RAG," consulting authoritative sources like IEEE Xplore, official library documentation, or textbooks. Students are prohibited from resolving conflicts by asking another AI model for the "correct" answer.

## Research Methodology

The study employed a mixed-methods approach combining quantitative psychometric instruments with qualitative behavioral analysis. This triangulation strategy enhances validity by capturing both self-reported cognitive capabilities and actual behavioral performance.

## Quantitative Analysis by Psychometric Instruments

### *Quantitative Instrument 1: Critical Thinking Scale (CTS)*

The Critical Thinking Scale (CTS) served as the primary quantitative measure of cognitive development. This instrument was adapted from two established sources: Facione's (1990) California Critical Thinking Disposition Inventory (CCTDI) and Ng et al.'s (2021) AI Literacy framework.

### *Quantitative Instrument 2: Complex Problem-Solving Assessment (CPS)*

The Complex Problem-Solving (CPS) Assessment represents a behavioral coding and scoring tool rather than a self-report instrument. This assessment was designed to capture actual performance in AI-assisted problem-solving contexts, addressing the limitation that self-report measures may not accurately reflect actual behavior. The CPS Assessment draws on two frameworks: (1) the OECD's (2014) PISA framework for assessing complex problem-solving, which emphasizes exploration, understanding, and planning in dynamic environments; and (2) Wing's (2006) computational thinking framework, which identifies key cognitive processes in technical problem-solving including decomposition, pattern recognition, abstraction, and algorithm design.

## Qualitative Behavioral Analysis

In addition to quantitative instruments, the study collected rich qualitative data.

### *Interaction Logs*

As described above, complete logs of student-AI interactions provided detailed behavioral data. Beyond quantitative coding, these logs were analyzed qualitatively to identify patterns, strategies, and changes in interaction styles.

### *Deliberation Logs*

Students submitted Deliberation Logs with each assignment, documenting their multi-model comparison, conflict identification, verification process, and reflections. These logs provided insight into students' deliberative reasoning and metacognitive awareness.

### *Reflection Journals*

Students maintained weekly reflection journals throughout the study, responding to prompts such as "What did you learn about AI capabilities and limitations this week?" and "How has your approach to using AI changed?" These journals captured students' evolving understanding and attitudes.

### *Focus Group Interviews*

At the conclusion of the study, three focus group interviews were conducted (10 students per group) to gather in-depth qualitative feedback on the MDGAIF experience, challenges encountered, and perceived benefits.

## Research Results

### Quantitative Findings

- CTS scores increased from 24.35 to 33.95 ($p < .001, d = 2.65$).
- Information filtering improved by 75%.
- Logical inference improved by 44%.
- CPS showed 237% growth, with review and identification improving 300%.

### Qualitative Findings

- Students shifted from vague, single prompts to multi-stage deliberative workflows.
- Error detection rates rose to 85% (logic) and 92% (factual).
- Reflection journals revealed increased awareness of AI limitations and ethical responsibility.

## Conclusion

The introduction of generative AI into higher education marks a significant milestone in educational technology evolution. Previously, most tools focused on making learning more efficient or convenient, but generative AI has the potential to fundamentally transform what "learning" means. Consider what if students could instantly come up with flawless answers without thinking. Then, what would become education's fundamental goal: to "develop critical thinking skills"? However, this study indicates that with the right teaching strategies, AI can support and enhance human thinking rather than replace it. Positioning AI as a tool for reflective inquiry, rather than a substitute for thought, facilitates the cultivation of critical thinking. This is achieved through a systematic process of student-led content verification and the establishment of clear accountability for the resulting contents. In response to this paradigm shift, MDGAIF offers a possible solution. It is grounded in solid theory, provides clear instructional steps, and has been thoroughly validated through research. While it is not a universal remedy, it offers a practical framework for educators aiming to integrate AI into teaching. This is without compromising academic quality and integrity. As AI capabilities advance unpredictably, MDGAIF principles-critical assessment, cross-verification, self-awareness, and ethical responsibility—will become increasingly vital. Regardless of whether teaching materials originated from humans or AI, teaching students to quickly evaluate information authenticity will inevitably become a key component of higher education. This study demonstrates that education, by keeping pace with AI and harnessing its power for positive development, is not only a contemporary trend but also an attainable goal. It gives us hope—higher education can undoubtedly adapt to the AI era and continue to fulfill its irreplaceable role of "inspiring the human intelligence."

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that AI technologies were used in the preparation of this manuscript. Specifically, Gemini Nanbanana was used to visually render Figure 1 based on the author's original conceptual design. Additionally, Gemini AI Tools was used for language refinement and proofreading to ensure clarity and grammatical accuracy. The ideas, research procedures, data analysis, and the core deliberative logic of the MDGAIF framework remain the original work of the author.

## References

Anthropic. (2025). *Claude 3.7 Sonnet: The first hybrid reasoning model*. Anthropic News. https://www.anthropic.com/news/claude-3-7-sonnet

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203771587

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. California Academic Press. https://eric.ed.gov/?id=ED315423

Mollick, E. R., & Mollick, L. (2023). *Using AI to implement effective teaching strategies in classrooms: Five strategies, including prompts*. Wharton School Working Paper. SSRN. https://ssrn.com/abstract=4391243

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041

OECD. (2014). *PISA 2012 results: Creative problem solving: Students' skills in tackling real-life problems (Volume V)*. OECD Publishing. https://doi.org/10.1787/9789264208070-en

OpenAI. (2024). *Technical report: Enhancing reasoning and verification in o1 models*. OpenAI Blog. https://openai.com/index/learning-to-reason-with-llms/

UNESCO. (2025). *AI and the future of education: Disruptions, dilemmas and directions*. UNESCO Publishing. https://unesdoc.unesco.org/ark:/48223/pf0000395236

Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33–35. https://doi.org/10.1145/1118178.1118215

Zhang, Y., Wang, L., Chen, C., & Liu, J. (2025). *Reasoning large language model errors arise from hallucinating critical problem features*. arXiv. https://doi.org/10.48550/arXiv.2505.12151

# Building an AI Teaching Assistant With RAG: Implementation and Performance Scoring Framework

Alexandre Moudjeb, Karlsruhe University of Applied Sciences, Germany
Fahmi Bellalouna, Karlsruhe University of Applied Sciences, Germany

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

AI-based assistants are increasingly used in educational contexts, yet large language models (LLMs) often produce inaccurate or hallucinated information, limiting their reliability for learning. This study investigates how Retrieval-Augmented Generation (RAG) can enhance the factuality and coherence of AI responses without additional model fine-tuning. We implement a multimodal educational assistant using RAG, capable of integrating textual documents, code, and images, and evaluating it on a dataset of 2,000 queries sourced from real student interactions and standardized educational exercises. Performance is measured using two complementary metrics: scientific factuality, assessing the accuracy of generated content against verified sources, and multimodal coherence, evaluating consistency across text, code, and images. A novel scoring method based on Jensen-Shannon divergence and temperature softmax is introduced, enabling automatic and dynamic evaluation of the system's responses. Validation against a set of human-annotated queries confirms the robustness of the scoring framework. Experimental results show that integrating RAG improves response accuracy by 9% and increases overall performance, demonstrating that RAG is a decisive factor in delivering reliable educational content. The multimodal approach further enhances the assistant's ability to address diverse query types, from textual explanations to code examples and visual illustrations. These findings suggest that RAG-based AI assistants can provide more trustworthy and pedagogically effective support, potentially improving student engagement and learning outcomes.

*Keywords:* large language models, RAG, education, scoring, performance

iafor
The International Academic Forum
www.iafor.org

## Introduction

In Germany and everywhere in the world, AI-based learning assistants are increasingly part of learning contexts, from helping learners ask questions to clarify misunderstandings or solve problems. Large Language Models (LLMs) have significantly enhanced the quality of natural language-based interactions; they still suffer from important limitations. They can produce plausible but wrong or unsupported answers like a hallucination phenomenon (Ji et al., 2023; Maynez et al., 2020). Such limitations indicate that generative models essentially rely on static knowledge acquired during training, without intrinsic mechanisms by which the model can check its responses against external sources.

These limitations have been addressed through a variety of approaches, including domain-specific adaptation or fine-tuning. However, these methods usually require large amounts of annotated data and are computationally expensive, besides being difficult to keep up to date in evolving domains. In such a context, Retrieval-Augmented Generation has emerged as a paradigm that appears very effective (Lewis et al., 2020). It combines text generation by an LLM with an information retrieval phase that resorts to relevant external sources at inference time. By combining retrieval and generation techniques, RAG thus allows answers grounded in explicit documents, reducing the model's reliance on internal memory, and improving factual accuracy without expensive fine-tuning.

In the context of education, RAG is particularly interesting because it allows for integration with various sources such as textbooks, course materials, code snippets, images and can answer complex questions by explicitly leveraging these resources (Holmes et al., 2019; Radford et al., 2021). Despite this interest, few studies offer a systematic analysis of multimodal RAG architectures for educational assistants, or an evaluation framework adapted to these systems (Lewis et al., 2020; Radford et al., 2021).

Evaluating such multimodal assistants remains a major challenge. Standard automatic metrics, designed for unimodal tasks, often fail to capture essential criteria such as scientific correctness, conceptual coherence, or alignment across modalities such as text, image, or code (Maynez et al., 2020). Human evaluations, though instructive, are expensive, subjective, and hard to generalize on a large scale (Holmes et al., 2019). There is, therefore, an increasing need for automated yet rigorous evaluation frameworks that not only measure response accuracy but also multimodal coherence (Radford et al., 2021). That is the challenge that this paper aims to address by proposing a structured framework to design, implement, and evaluate a multimodal educational assistant based on RAG. This paper presents an implementation developed in FastAPI, in a modular software architecture that conforms to the principles of object-oriented programming; it embeds a multimodal pipeline that can process textual documents, code snippets, and images. Besides implementation, this paper proposes a light, but rigorous scoring framework based on two complementary dimensions: scientific factuality and multimodal coherence. This approach allows the analysis of the respective and combined impact of dynamically retrieved documents and those indexed in the RAG knowledge base. Experimental results, with state-of-the-art multimodal models, show that the integration of RAG significantly improves response accuracy and overall performance, underlining the interest of such an approach for deploying reliable and effective educational assistants.

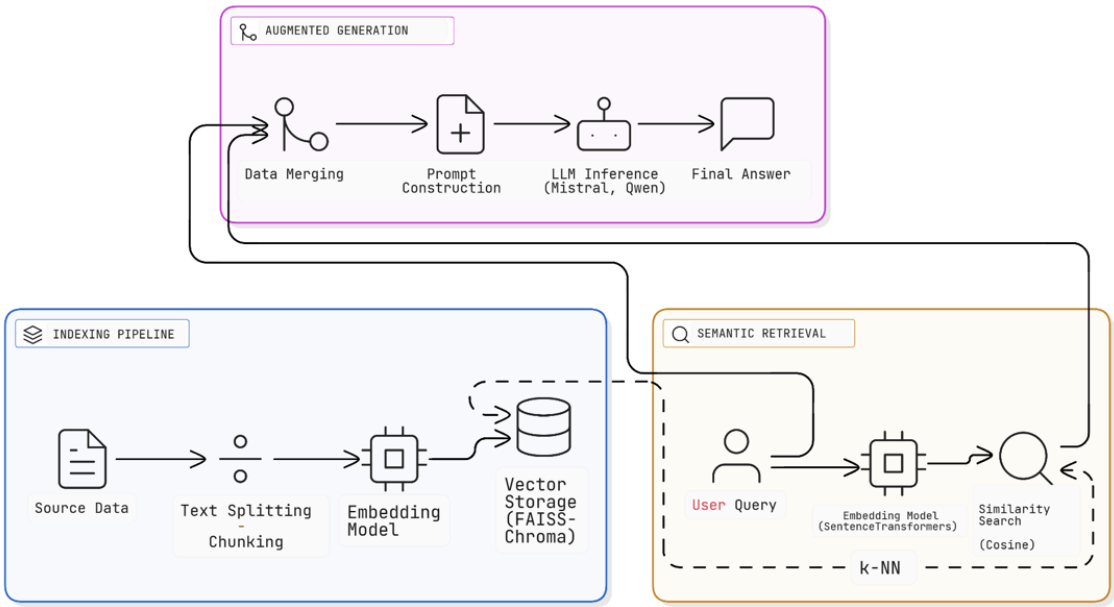
## Background and Related Works

### Retrieval Augmented Generation for Educational Assistants

In the context of AI-automated pedagogy, RAG is a powerful tool for increasing factuality. Indeed, finely trained models can provide correct answers across a wide range of domains to the extent that they have been trained, although certain niches or precise details may remain problematic. Consequently, for more specific knowledge or fine details of reasoning in problem solving, the model alone may not have been previously trained, and it may prove difficult to obtain a reliably generated answer, which can increase the risk of hallucination (Lewis et al., 2020).

Nowadays, there are several ways to implement an RAG system. The most common system is the use of a vector database (FAISS, Chroma) containing documents vectorized beforehand by an embedding model (e.g., a model from Sentence-Transformers). This database becomes a limited source, whose reliability depends on the quality of the indexed information, and through various processes (development of utility functions), by defining the Euclidean dot product between the components of the database vectors and those of the query, it is possible to select the vectors closest to the query and, ultimately, retrieve the important passages that the large language model can reinterpret to generate its answer (Abo El Enen et al., 2025). It would therefore be consistent to use this tool for pedagogical purposes.

**Figure 1**
*Architectural Diagram*




*Note.* Indexing of documents via FAISS, the vectorization of documents and the query, as well as the post-query processing leading to the generation of a response.

The use of vector database technology in the RAG process is called Dense Retrieval (Karpukhin et al., 2020; Lewis et al., 2020). Indeed, vectorization followed by comparison using similarity computation via L2 distance and cosine similarity allows for a precise quantification of the proximity between queries and passages, which makes it possible to characterize meaning and perform semantic search (Devlin et al., 2019; Karpukhin et al., 2020). What may be interesting is to couple this system with an algorithm such as BM25, which allows counting the frequency of exact words and can give weight to passages

containing a particular word. This technology is called Sparse Retrieval (Robertson & Zaragoza, 2009).

## Informations Regarding FastAPI Application Development

FastAPI is designed as an efficient framework for backend development with applications in data-intensive applications as well as artificial intelligence projects (Ji et al., 2023; Maynez et al., 2020). Founded on Starlette for handling requests asynchronously and Pydantic for structuring data, it is built using Python's type annotations system that fosters reliability, automatic input validation, and the generation of OpenAPI documents dynamically. With capabilities tailored for managing multiple requests asynchronously, it can successfully process multiple document uploads, vector search, or large-scale language model evaluations at an efficient rate (Bednarz & Milosz, 2025; Chen, 2023).

In retrieval augmented generation (RAG) models, FastAPI manages these tasks: preprocessing of documents, embedding computation, vector indexing, and response generation. The modularity of FastAPI along with dependency injection enables experiments to be performed as well as replacing tasks such as embedding models without affecting other tasks (Brown et al., 2025; Chen, 2023; Gao et al., 2023).

It also allows file upload, response streaming, and multimodal data, which is necessary when working with educational assistants that process texts, code, and scientific documents. Response streaming improves interaction in conversational systems. With the help of ASGI-based solutions, FastAPI provides an efficient and reproducible method used in the development of superior artificial intelligence applications (Bednarz & Milosz, 2025; Chen, 2023; Haque, 2025).

## Current Scoring Methods and Performance Evaluation Metrics

It is pertinent to mention that multimodal assistants are assessed using a set of evaluation measures to judge alignment, coherence, as well as reasoning across various modalities like text, images, or more. Among various evaluation measures used to judge alignment between multimodal pairs like an image and text, CLIPScore facilitates evaluating semantic coherence between an image and text based on the cosine similarity between them in an embedding space learned using contrastive learning (Hessel et al., 2021). Though this is helpful in judging alignment across various multimodal pairs, this measure is restricted to representational similarity only between the multimodal pairs in question (Radford et al., 2021).

On the same line, the image and text-based retrieval scores, like Recall@K, Mean Reciprocal Rank (MRR), and scores like EM1 and EM2, are used largely for the assessment of the quality of the ranks and the accuracy level of the respective responses in the context of searching and/or questioning and answering processes (Antol et al., 2015; Manning et al., 2008). EM1 calculates the exact level of the produced answer retrieved against the reference answer, whereas EM2 accounts for the tolerant form variations in the respective correct responses to be retrieved. Additionally, semantic similarity scores like SS1 and SS2 can be used for measuring the semantic similarity levels between the answer and the reference, where SS2 considers the overall weights of the segments or the modalities (Antol et al., 2015; Yu et al., 2023).

For those tasks involving complex reasoning with multimodal information, such as Visual Question Answering, ScienceQA, or MMMU challenges, the tasks are assessed using accuracy

measures to calculate the ratio of correct answers on labeled examples (Clark et al., 2022). Furthermore, recent work suggests the utilization of language models acting as independent assessors to judgeOPEN ENDED responses to multimodal tasks with respect to notions of coherence, relevance, and multimodal response faithfulness.

While the strategies are highly promising, some level of subjectivity and continued reliance upon the evaluator model are still incorporated. Another method that has been attempted in the calculations of dependence among modalities on a theoretical level is the Mutual Information (MI) metric, which estimates the difference between the joint distribution of the modalities and the product of their marginal distributions (Cover & Thomas, 2006). Nevertheless, despite the robust theoretical background of the measure, there exist several limitations in applying the MI metric in a multimodal framework. The measure has no upper boundary and faces the problem of inaccurate estimates in high-dimensional representation spaces, besides being statistically noisy (Belghazi et al., 2018). Besides, the extension of this measure to large modalities remains a difficult task, making it less preferred for the task of multimodal evaluation in large scenarios.

## Global System Design and Architecture

### Functional Design of the Assistant and Pedagogical Objectives

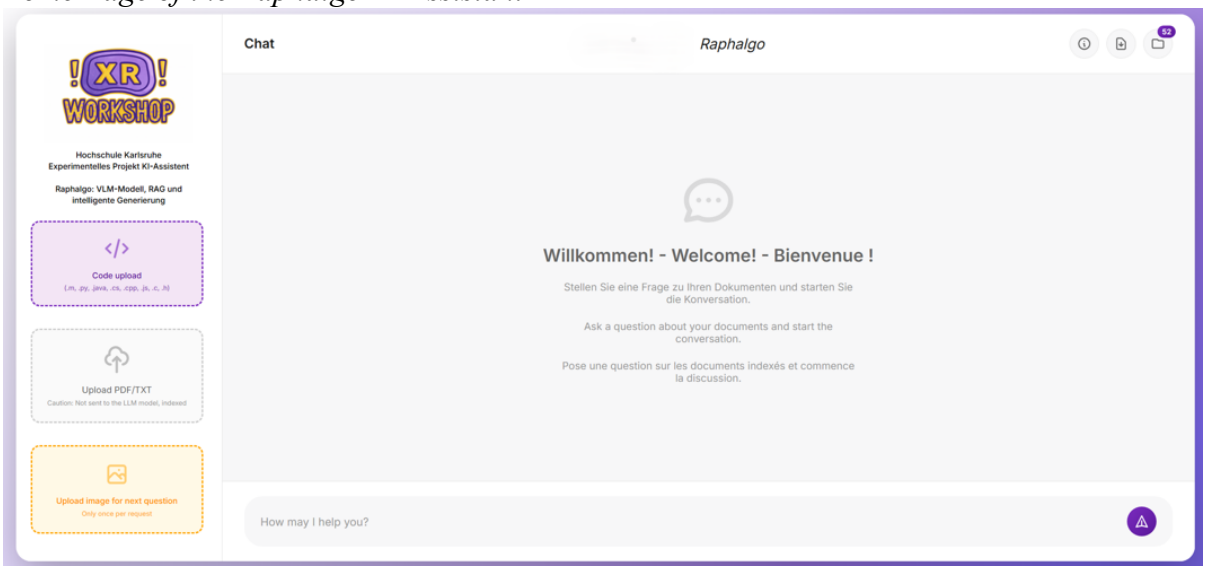
The proposed pedagogical assistant is designed as a multimodal educational system capable of answering questions in natural language, explaining technical concepts, analyzing source code, and interpreting visual statements. The assistant is restricted to an academic context in which factual accuracy, source traceability, and pedagogical coherence constitute, in our view, requirements to limit the hallucination phenomena observed in language generation models (Ji et al., 2023; Maynez et al., 2020).

The user can upload PDF or TXT documents, which constitute the assistant's knowledge base. These documents are never transmitted directly to the language model: they are indexed locally to serve as support for the information retrieval mechanism. This architectural choice is consistent with the principles of retrieval-augmented generation architecture, which aim to anchor generation in explicit and verifiable sources (Brown et al., 2025; Gao et al., 2023; Lewis et al., 2020). It also guarantees better data confidentiality and enhanced control over the content used during the generation of responses (Haque, 2025).

The assistant also supports the upload of code files (MATLAB, Python, C/C++, Java, etc.), which are stored at the user session level. This design enables iterative code analysis and pedagogical continuity across multiple interactions, in line with the educational uses of AI assistants specialized in engineering and computer science (Holmes et al., 2019). Finally, images can be associated with a single query to process exercises or problems containing visual elements, aligning with multimodal approaches studied in the fields of Visual Question Answering and AI-assisted scientific learning (Antol et al., 2015; Clark et al., 2022).

A bounded conversation history is maintained to ensure contextual coherence while limiting the size of the requests sent to the model. The assistant systematically adapts the language of its responses to that of the user, promoting multilingual use in an international educational context.

**Figure 2**
*Home Page of the Raphalgo AI Assistant*




*Note.* Based on Mistral (to the left, the drag-and-drop areas; at the bottom, the search bar; and at the top right, the HTML export, the document registry, and an information tab).

## RAG-Based Architecture and Inference Pipeline

The architecture of the assistant is based on a Retrieval-Augmented Generation (RAG) approach combining a dense vector search engine and a large language model. This architecture, initially introduced for knowledge-intensive tasks, aims to improve the factuality and robustness of the responses generated (Gao et al., 2023; Lewis et al., 2020).

The uploaded documents are first preprocessed through text extraction (PDF parsing or direct reading), then segmented into overlapping textual fragments to preserve semantic continuity. Each fragment is then transformed into a dense vector using a sentence-transformer–type embedding model based on MiniLM (all-MiniLM-L6-v2) (Reimers & Gurevych, 2019; Wang et al., 2020). These approaches rely on pre-trained bidirectional transformer architectures, widely validated for the semantic representation of sentences and passages (Devlin et al., 2019; Goodfellow et al., 2016).

The generated vectors are stored in a FAISS (Facebook AI Similarity Search) index, designed for large-scale vector similarity search, and widely used in dense retrieval systems (Johnson et al., 2019). The associated metadata, including the document name and the textual content of the fragments, are preserved to ensure source traceability, a key point in educational and scientific systems.

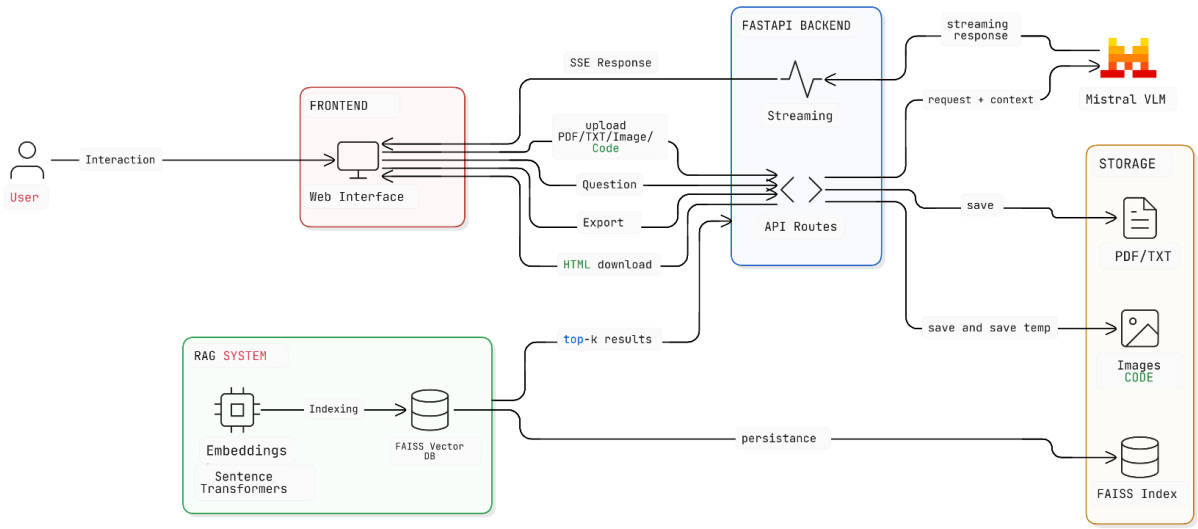
When a user query is received, it is vectorized and compared against the FAISS index to extract the k most relevant fragments, according to a dense semantic proximity measure (Karpukhin et al., 2020). These fragments are concatenated to form a documentary context injected into the prompt sent to the language model. In the absence of relevant results in the index, a fallback mechanism extracts a limited number of fragments directly from the available documents.

The FAISS index is persistent and automatically synchronized with the document directory via a process executed in the background. This approach enables the dynamic addition or removal of content without service interruption. It also makes it possible to overcome the limitations of

traditional lexical search methods, such as BM25, which rely primarily on exact term matching (Manning et al., 2008; Robertson & Zaragoza, 2009). In scientific fields such as physics or control theory, this ability to capture conceptual proximity is essential for linking physical principles to their mathematical formulations, even when terminology varies across pedagogical materials. For indexing, the text is split into fragments of 600 tokens with an overlap of 150 tokens, ensuring semantic continuity across the 32,678 generated vectors and limiting context loss in complex scientific demonstrations.

The entire system is orchestrated by a FastAPI API, which manages document ingestion, user queries, and the streaming delivery of responses. The choice of FastAPI is based on its performance and low latency in intensive application contexts (Bednarz & Milosz, 2025; Chen, 2023). The language model is executed locally and queried via a generation interface allowing the progressive return of tokens.

**Figure 3**
*Global Architecture of the RAG AI Assistant*




## Prompt Orchestration, Multimodality, and Output Control

The behavior of the assistant is driven by a prompt orchestration strategy integrating several components: system instructions, documentary context derived from the RAG, conversation history, and optional multimodal inputs. This structuring aligns with recent work showing that explicit prompt design improves reasoning quality and reduces unfounded generation errors (Wei et al., 2022).

A fixed system instruction explicitly defines the pedagogical role of the assistant, enforces linguistic consistency, and specifies strict formatting rules, particularly for mathematical formulas and code blocks. When a code file is present, it is included in the prompt and treated as a central reference for the analysis and explanations provided.

In the case of multimodal queries, the uploaded images are encoded and attached to the request to enable joint text–image reasoning, in accordance with work on multimodal evaluation and reasoning in large-scale models (Antol et al., 2015; Hessel et al., 2021; Yu et al., 2023). The generation parameters are deliberately conservative, notably using a low temperature, to limit degeneration of the generated text and to improve factual stability and reproducibility of the responses (Holtzman et al., 2020).

The responses are produced in enriched Markdown, compatible with LaTeX rendering of mathematical expressions and subsequent export of conversations. Finally, information related to the retrieved context, the modalities used, and the dialogue history is logged, providing the elements necessary for the quantitative performance evaluation presented in Section 5.

**Figure 4**
*HTML Preview of the Report Generated by the User, Using Markdown and LaTeX Formatting*

## Scoring Method

### Scientific Factuality Score and RAG Precision

Let us consider a panel of n numbered queries. For a query i, we denote by $T_i$ the expected rigor corresponding to the degree of expectation assigned to the assistant, and by $\{T_i\}_{i\in[1,n]}$ the associated family. We proceed similarly for the estimated values of the assistant's responses, which we denote by $M_i$ and $\{M_i\}_{i\in[1,n]}$ the associated family. Each of these $T_i$, $M_i$ given as a percentage, is defined on a scale from 0 to 100%, with 1 if the assistant answers the question rigorously and completely.

Thus, one can deduce an average divergence Δ of scientific factuality given by the formula:

$$\Delta = \frac{1}{n} \sum_{i \in [1,n]} |T_i - M_i| \tag{1}$$

and an associated score S given by S = 1 – Δ.

For a RAG precision score, a simple formula is used by assigning each $T_i$ to the number of sources expected/that should be retrieved for a query i, and $M_i$ to the number of sources retrieved among the expected sources for this same query i. Note that in the case of omission, if a source is retrieved that was not anticipated in $T_i$, if it is coherent, it will trigger a +1 for both $T_i$ and $M_i$, namely:

$$\Delta = \frac{1}{n} \sum_{i \in [1,n]} \frac{M_i}{T_i} \tag{2}$$

Note that this divergence formula only makes sense if there exists a document dealing with the question. It should be noted that this score does not consider the ranking of the retrieved information, which is a minor issue if the top-k retrieved items are sufficient on their own to produce a complete answer.

**Multimodal Consistency Score**

The objective of this scoring is to evaluate the internal consistency of the reasoning of a scientific assistant when it relies on heterogeneous sources (text, code, documents). By comparing the quality profiles obtained from responses conditioned on a single source, the score measures whether these sources support a compatible scientific interpretation or produce divergent reasoning.

It does not aim to judge the absolute correctness of a response, but to quantify the structural alignment of the contributions of each source to the overall reasoning.

We therefore explain the method for obtaining this score. In a first step, a query q is submitted to the assistant with a particular prompting strategy; that is, the same question is asked, but the assistant is instructed to respond using only one type of source at a time (either PDF documents, text, or code).

**Table 1**
*Adapted Prompting*

| Prompt 1 – Text only |
|---|
| Answer the following question using only text excerpts from the provided TXT documents. You may use an unlimited number of words, as long as the reasoning is, in your judgment, sufficiently thorough and complete.<br><br>If you cannot find information that you consider coherent and relevant to the query, return "not retrieved or unsufficient context".<br>Context: "TEXT documents". |
| Prompt 2 – PDF only |
| Answer the following question using only text excerpts from the provided PDF documents. You may use an unlimited number of words, as long as the reasoning is, in your judgment, sufficiently thorough and complete.<br><br>If you cannot find information that you consider coherent and relevant to the query, return "not retrieved or unsufficient context".<br>Context: "PDF documents". |

| Prompt 3 – Code only |
| --- |
| Answer the following question using only text excerpts from the provided Code documents. You may use an unlimited number of words, as long as the reasoning is, in your judgment, sufficiently thorough and complete.<br><br>If you cannot find information that you consider coherent and relevant to the query, return "not retrieved or unsufficient context".<br>Context: ".py, .m, .cs, .java, .c, .h documents". |

From these generated responses, one can assess (either manually or via an LLM-judge) a certain score according to a few common criteria B (e.g., scientific clarity, logical coherence, absence of hallucinations, structuring, possible numerical applications, etc.), such that for a query q, we obtain the score of modality m, denoted $s_{q,m}$ composed of:

$$s_{q,m} = (s_{q,m}(1), \dots, s_{q,m}(B)) \text{ with } s_{q,m} \in [\![0,1]\!]^B \tag{3}$$

To this per-query score, one can associate a statistical distribution $L_{q,m}$, where the score vector is interpreted as an empirical distribution over the criteria:

$$L_{q,m} = (L_{q,m}(1), \dots, L_{q,m}(B)) \text{ with } L_{q,m} \in [\![0,1]\!]^B \text{ such as } L_{q,m}(b) = \frac{s_{q,m}(b) + \varepsilon}{\sum_{b \in [\![0,B]\!]} s_{q,m}(b) + \varepsilon}$$

$\varepsilon = 10^{-12}$ (Holtzman et al., 2020), $b \in [\![1,B]\!]$ (4)

Through the normalization imposed by the distribution, information related to exactness is lost; nevertheless, this is not problematic since the goal is to obtain a score related to the coherence between several statistical distributions. There currently exist several functions for measuring coherence between multiple distributions, including the Jensen–Shannon divergence, which in our case allows comparison of the relative structure of reasoning qualities induced by each source involved.

The objective is to compare statistical distributions associated with different modalities, each distribution representing the way a given source supports reasoning according to several criteria. Since there is no natural reference distribution nor dominant modality, the comparison must be symmetrical and independent of the order of the sources. As divergence between distributions is defined for two distributions at a time, multimodal comparison naturally reduces to the study of all binary comparisons between modalities. The triangular sum then aggregates these pairwise divergences exhaustively and non-redundantly, providing a global measure of the intermodal coherence of the reasoning.

Ultimately, one first computes the mean of the pairwise Jensen–Shannon divergences between modalities for a given query and then averages these divergences over the number of queries performed. Considering Q as the number of queries asked and M as the number of modalities, the expression is:

$$\Delta = \frac{2}{M(M-1)Q} \sum_{q \in [1,Q]} \sum_{1 \le i < j \le M} D_{JS}{}^*(L_{q,i} \parallel L_{q,j}) \tag{5}$$

Where $D_{JS}{}^*$ denotes the normalized Jensen–Shannon divergence function.

In terms of algorithmic complexity, we have $O(M^2Q)$. For a limited number of modalities (e.g., three modalities: code, text, and PDF documents), the computation of the sum can be performed quickly.

An associated score is defined using the previous formula S = 1 – Δ.

**Multimodal Performance Indicator**

From these scores, one can define a global performance indicator measuring how coherent and reliable a scientifically oriented AI assistant could be. We perform an aggregation of our scores. This represents a structural measure of trust.

Let $\boldsymbol{C}$ denote the set of score components (RAG, Multimodal Consistency Score, scientific scores, latency, etc.).

We formalize this indicator by computing a weighted sum of the scores with carefully chosen positive weights $\boldsymbol{w_c}$, whose sum equals 1. Thus, denoting the indicator by IP, we have:

$$\boldsymbol{IP} = \sum_{\boldsymbol{c} \in \boldsymbol{C}} \boldsymbol{w_c S_c} \tag{6}$$

The experimenter has full freedom in choosing the weights $\boldsymbol{w_c}$. We propose, for each weight, the use of the softmax formula applied to the divergences associated with the score, so that for each weight we have:

$$w_c = \frac{\exp\left(-\lambda \Delta_c\right)}{\sum_{j \in C} \exp\left(-\lambda \Delta_j\right)} \tag{7}$$

with λ the sensitivity parameter, which we define as:

$$\lambda = \frac{1}{med\left(\{\Delta_j\}_{j \in C}\right) + \varepsilon} \tag{8}$$

The sensitivity parameter allows us to adjust the weights according to our tolerance for errors. This expression for λ is self-adaptive, and the inverse of the median is a robust estimator of the scale of the divergences, allowing the parameter to be normalized by the natural scale of the observed phenomenon (Huber, 1981).

Nevertheless, if the experimenter wishes to discriminate errors by penalizing the weights of scores whose divergence is maximal, they can choose this λ.

Thus, assuming that each score can be written in the form $\boldsymbol{S_c} = \boldsymbol{1} - \boldsymbol{\Delta_c}$, the general form of the performance indicator is:

$$IP = \sum_{c \in C} (1 - \Delta_c) \left( \frac{\exp\left(-\lambda \Delta_c\right)}{\sum_{j \in C} \exp\left(-\lambda \Delta_j\right)} \right) \tag{9}$$

## Results

### Evaluation Method

In the first step, we are going to calculate standard metrics of precision, recall, and ranking scores to globally assess the performance of the assistant with and without RAG, then compare its added value in the modeling by using the proposed scoring method.

### RAG Impact

To evaluate the performance of our AI assistant, we will first carry out an initial measurement of the effectiveness of RAG by computing the standard metrics Mean Reciprocal Rank (MRR), Global Precision (P), Global Recall (R), and top-6 Precision/Recall (P@6/R@6) on a panel of 2,000 queries covering the following scientific topics (thermodynamics, mechanism dynamics, electronics, mathematics, computer programming, optics) (Manning et al., 2008).

**Table 2**
*Evaluation Metric Values*

| Metrics | MRR | P | R | P@6 | R@6 |
|---|---|---|---|---|---|
| RAG | 0,756 | 0,772 | 0,827 | 0,723 | 0,781 |
| NO RAG | / | 0,689 | 0,694 | / | / |

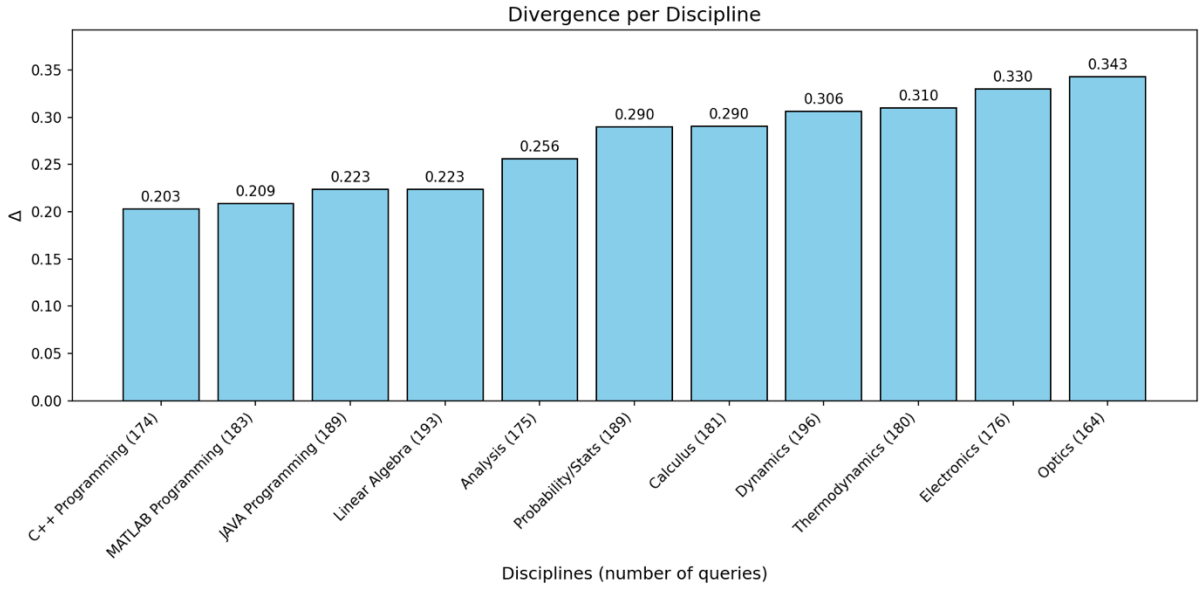
The model achieves an R@6 of 0.781 and a P@6 of 0.723, indicating that most relevant items are retrieved among the top six results while maintaining a high level of precision. Combined with an MRR of 0.756, these performances show that relevant items are generally ranked in the very first position or very close to the top of the list, reflecting the model's effective ability to correctly order results according to their relevance. It is also observed that the RAG system increases precision by 9% and global recall by 12%, which implies a performance improvement of 10.5% according to the F1-score. The RAG system is effective. MRR and Recall are not defined in the NO-RAG setting, as no retrieval mechanism is involved.

### Scoring Method

For our scoring method, we will establish a scientific factuality score for each of the considered disciplines. We will also consider our overall RAG score. For the modality score, we will consider the provided text, code, images, and PDF documents. Next, we will compute the associated parameters and the corresponding performance indicator (IP).

We plot the histogram of the divergences $\Delta$ for the tested academic subjects, using the number of queries derived from university-level exercises.

**Figure 5**
*Histogram of Divergence per Discipline*




We observe that the most satisfactory responses correspond to programming, as they minimize the divergence. This may be due to several reasons, including the choice of the embedding model, which ignores image files that are more prevalent in physics courses. The assistant struggles more with the physical sciences, which is consistent with the experimenter's observations. These values are averaged to facilitate the analysis.

On the same set of 2,000 queries, the RAG system made it possible to retrieve, on average, 74.7% of the expected sources. For the multimodality score, we considered four criteria: scientific clarity, logical coherence, absence of hallucinations, and correctness of numerical applications. For this purpose, we applied the prompting shown in Fig. 5 to obtain five responses per query (code, text, PDF, image, and intermodal response) for the joint evaluation of $\Delta_S$ and $\Delta_{RAG}$).

**Table 3**
*Study Values for the Performance Indicator*

| | M | Q | $\Delta_S$ | $\Delta_{RAG}$ | $\Delta_M$ |
|---|---|---|---|---|---|
| Values | 4 | 2000 | 0,271 | 0,253 | 0,324 |

We obtain a modality divergence of 0.324. This means that the AI assistant can synchronize the different sources to promote a structured and coherent response across these sources with an alignment of 67.6%. The reasoning is generally well aligned.

**Table 4**
*Performance Indicator Measurements*

| | $\Delta_S$ | $\Delta_{RAG}$ | $\Delta_M$ | $\lambda$ | $w_S$ | $w_{RAG}$ | $w_M$ | IP |
|---|---|---|---|---|---|---|---|---|
| Values | 0,271 | 0,253 | 0,324 | 3,690 | 0,346 | 0,370 | 0,284 | 0,720 |

Our scoring method evaluates our RAG-based AI assistant with a score of 7.2/10, which at first glance is relatively satisfactory. The model used may (or may not) discriminate strongly between high values, which can influence the final score by diverging from the original intent

of the scoring scheme. We observe that the model assigns a relatively balanced importance to all scores.

## Conclusion

This work shows that integrating RAG clearly improves the reliability of an educational assistant by reducing hallucinations and increasing the accuracy of responses. The proposed architecture functions effectively and remains adaptable to different types of data.

The scoring method provides a more relevant evaluation, particularly regarding multimodal coherence, although it remains imperfect and strongly dependent on the experimenter. The results confirm a real performance gain, but also highlight limitations on certain types of content, especially in physics or when dealing with visual data.

RAG therefore appears to be an effective solution, but one that strongly depends on the quality of the sources and the technical choices made.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that Grammarly, an AI-assisted writing software, was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The authors further declare that, apart from Grammarly, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Abo El Enen, M., Saad, S., & Nazmy, T. (2025). A survey on retrieval-augmented generation (RAG) models for healthcare applications. *Neural Computing and Applications, 37*(33), 28191–28267. https://doi.org/10.1007/s00521-025-11666-9

Antol, A., Agrawal, A., Lu, J., Mitchell, M., Batra, D., Zitnick, C. L., & Parikh, D. (2015). VQA: Visual question answering. In *Proceedings of the IEEE International Conference on Computer Vision* (ICCV) (pp. 2425–2433). https://doi.org/10.1109/ICCV.2015.279

Bednarz, B., & Milosz, M. (2025). Benchmarking the performance of Python web frameworks. *Journal of Computer Sciences Institute, 36*, 336–341. https://doi.org/10.35784/jcsi.7738

Belghazi, M. I., Baratin, A., Rajeswar, S., Ozair, S., Bengio, Y., Courville, A., & Hjelm, R. D. (2018). Mutual information neural estimation. In *Proceedings of the International Conference on Machine Learning*, (Vol. 80, pp. 531–540). PMLR.

Brown, A., Roman, M., & Devereux, B. (2025). *A systematic literature review of retrieval-augmented generation: Techniques, metrics, and challenges*. arXiv. https://arxiv.org/abs/2508.06401

Chen, J. (2023). Model algorithm research based on Python FastAPI. *Frontiers in Science and Engineering*, *3*(9), 7–10. https://doi.org/10.54691/fse.v3i9.5591

Clark, P., Lu, P., Mishra, S., Xia, T., Qiu, L., Chang, K.-W., Zhu, S.-C., Tafjord, O., & Kalyan, A. (2022). Learn to explain: Multimodal reasoning via thought chains for science question answering. In *Advances in Neural Information Processing Systems, 35*. https://arxiv.org/abs/2209.09513

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed.). Wiley.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (Vol. 1, pp. 4171–4186). Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1423

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2023). *Retrieval-augmented generation for large language models: A survey*. arXiv. https://arxiv.org/abs/2312.10997

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Haque, I. C. R. (2025). Implementation of RAG and LLM for a document-based chatbot system. *Journal of Electronics Technology Exploration, 3*(1), 19–23. https://doi.org/10.52465/joetex.v3i1.588

Hessel, J., Holtzman, A., Forbes, Y., Bras, R., & Choi, Y. (2021). CLIPScore: A reference-free evaluation metric for image captioning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (pp. 7514–7528). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.595

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. OECD Publishing.

Holtzman, A., Buys, J., Du, L., Forbes, M., & Choi, Y. (2020). The curious case of neural text degeneration. In *Proceedings of the International Conference on Learning Representations*. https://arxiv.org/abs/1904.09751

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, *55*(12), 1–38. https://doi.org/10.1145/3571730

Johnson, J., Douze, M., & Jégou, H. (2019). Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data, 7*(3), 535–547. https://doi.org/10.1109/TBDATA.2019.2921572

Karpukhin, V., Oguz, B., Min, S., Lewis, P., Wu, L., Edunov, S., Chen, D., & Yih, W.-t. (2020). Dense passage retrieval for open-domain question answering. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 6769–6781). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-main.550

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems, 33* (pp. 9459–9474).

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.

Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). On faithfulness and factuality in abstractive summarization. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 1906–1919). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.173

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning transferable visual models from natural language supervision. In *Proceedings of the 38th International Conference on Machine Learning* (Vol. 139, pp. 8748–8763). PMLR.

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing* (pp. 3982–3992). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-1410

Robertson, S., & Zaragoza, H. (2009). The probabilistic relevance framework: BM25 and beyond. *Foundations and Trends in Information Retrieval*, *3*(4), 333–389. https://doi.org/10.1561/1500000019

Wang, W., Wei, F., Dong, L., Bao, H., Yang, N., & Zhou, M. (2020). MiniLM: Deep self-attention distillation for labeled and unlabeled pre-trained language models. *Advances in Neural Information Processing Systems, 33*, 12941–12952.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems, 35*.

Yu, F., Yang, Z., Li, L., Wang, J., Lin, K., Liu, Z., Wang, X., & Wang, L. (2023). *MM-Vet: Evaluating large multimodal models for integrated capabilities*. arXiv. https://arxiv.org/abs/2308.02490

# Instructional Leadership in Underperforming Schools: Evaluating the Intervention Strategies of School Management Teams and Subject Advisors

Sandile Xaba, Sol Plaatje University, South Africa
Richard Moloele, Sol Plaatje University, South Africa
Maseeeng Papashane, Sol Plaatje University, South Africa

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

The National Curriculum Statement (NCS) results serve as a primary benchmark for school functionality in South Africa. Despite existing accountability frameworks, persistent underperformance remains a challenge. This qualitative study investigated intervention strategies employed by School Management Teams (SMTs) and subject advisors to enhance Grade 12 performance. Set in the in the Dr Ruth Sekgomotsi Mompati District Municipality and grounded in instructional leadership theory, the research utilized purposive sampling to select three high schools that have underperformed in Grade 12 for over fifteen years. Data were collected through in-depth interviews with three teachers, three principals, and two district officials, then analysed via inductive thematic analysis. The findings revealed a multi-pronged intervention approach categorised into three core themes: (1) Psychosocial Support, through motivational programme aimed at learner resilience; (2) Intensive Pedagogical Support, via academic camps to address curriculum gaps; and (3) Institutional Oversight, through human resource capacitation and subject advisor support. While these strategies are central to district recovery plans, their efficacy is hampered by geographic disparities and resource constraints. The study recommends that the Department of Education decentralise resource allocation to support remote schools and transition subject advisors from compliance monitoring to active instructional coaching. Further research is required to evaluate the long-term effectiveness of these specific interventions in sustainable academic recovery.

*Keywords:* academic underperformance, instructional leadership, school management teams, school leadership, subject advisor

iafor
The International Academic Forum
www.iafor.org

## Introduction

The South African National Curriculum Statements (NCS) Grade 12 results serve as a benchmark for evaluating system standards and assessing functionality (Potterton, 2015). Even though Taylor (2021) warns these metrics do not necessarily indicate deep conceptual understanding, results continue as the basis for comparative assessments (Allais, 2014). Accountability is a focal point within the DBE, where stakeholders are held accountable for performance (Klinck et al., 2023). This is manifested through intervention strategies, including funded programmes like winter schools, spring classes, incubation camps, and Saturday classes (Mkhize & Otu, 2018; Mpiti & Wambu, 2023) designed to insulate learners against systemic shortcomings (West & Meier, 2020).

When schools underperform, subject advisors, teachers, and principals are called into account. Under Personnel Administration Measures (PAM), SMTs and district offices are entrusted with executing curriculum and designing programmes to enhance academic performance (Department of Basic Education [DBE], 2016; Ehren et al., 2020; Mthembu, 2025). Tabe (2023) argues factors such as teacher quality affect performance, meaning teachers are at the centre of academic success. According to the DBE (2020), subject advisors' responsibilities include monitoring and supporting teachers in classrooms and implementing curriculum changes (Department of Basic Education, 2023; Mthembu, 2025). Mouton and Malumbete (2023) agree with DBE (2013) that chief responsibilities are oversight and assistance with curriculum implementation.

When performance-based accountability models fail, SMTs and advisors find themselves holding accountability for results complicated by deficiencies outside their control (Heystek, 2015). Some districts face underperforming schools where performance has been below 35% or under 50% for consecutive years. The Dr RSM district was at the bottom of Grade 12 rankings in North-West province for over fifteen years (DBE, 2022a). This prolonged underperformance suggests a disjuncture between intervention and classroom reality. There is a lack of qualitative understanding regarding how leadership roles manifest in specific strategies targeting recovery in Dr RSM district. This study asks: What are the types of intervention strategies the SMTs, and subject advisors employ to enhance the matric academic performance in Dr. RSM education district?

## Literature Review

### School Management Teams

The SMT comprises the principal, deputy principal, and Heads of Department, and may co-opt senior teachers and heads of support departments (Klinck et al., 2023). Under principal leadership, it plays a role in ensuring success. Planning and execution through teachers is central to SMT responsibilities. Teachers carry out tasks involving content delivery to support learners and enhance performance (Klinck et al., 2023; Maja, 2016; Mthiyane, 2019; Subramoney, 2016).

SMTs shape academic performance. Proactive efforts in strategic planning tailored to foster quality have a positive impact on learner performance (Chabalala & Naidoo, 2021; Maifala, 2024; Ndimande, 2005; Yingwane, 2025). Schools lacking strategic planning mechanisms experience underachievement. Research highlights shortcomings like failure of SMTs to

execute enhancement programmes, inadequate involvement of parents, insufficient teacher motivation, and ineffective communication (Munsamy, 2023; Netshisumbewa, 2024).

Teachers' dedication is influenced by motivation, support, and resource provisioning from SMTs (Hompashe, 2024; Munsamy, 2023), impacting performance (Manyaka, 2025; Mosibudi, 2012; Ntseto, 2015). Additionally, SMTs encourage parents to participate in school affairs and keep them informed (DBE, 2016, 2022b).

**Subject Advisors**

Subject advisors, also referred to as superintendents, inspectors, field or subject specialists (Mbande-Kwayini et al., 2024), are

> specialist office-based educators in a district office or circuit office whose function is to facilitate curriculum implementation and improve the environment and process of learning and teaching by visiting schools, consulting with and advising school principals and teachers on curriculum matters. (DBE, 2013, p. 11)

District offices serve as the link between Provincial Education Departments (PEDs), educational institutions, and the public (DBE, 2013, 2022b, 2023).

According to the Policy on the Organisation, Roles and Responsibilities of Education Districts (DBE, 2013), the first role of a Subject Advisor is to monitor the curriculum and support implementation. When monitoring, they observe practices, check progress, and provide remedies (Mouton & Malumbete, 2023; Munyai, 2024). Their chief roles are monitoring and supporting curriculum delivery to ensure quality teaching (DBE, 2022b). They oversee curriculum, ensure teachers are developed, evaluate learner performance, monitor HoDs, ensure availability of materials, and conduct class visits with feedback (DBE, 2016; Govindasamy & Mestry, 2022). Subject advisors play an instructional role, conduct workshops, facilitate clusters, and visit schools (Bantwini & Diko, 2011; Bantwini & Moorosi, 2018; Shushu, 2023). District offices and subject advisors work closely with teachers (Mbande-Kwayini et al., 2024; Mthethwa, 2022).

**Interrelationship Between Subject Advisors and Schools**

The relationship between schools and districts is pivotal. Continuous communication and collaboration are imperative. DBE (2013) underscores district officials collaborating with principals, providing guidance and support. Disharmony or misalignment due to insufficient resources, lack of consultation, limited visibility, and responsiveness to change impacts learner achievement (Bantwini & Moorosi, 2018; Mouton & Malumbete, 2023; Munyai, 2024; Netshisumbewa, 2024).

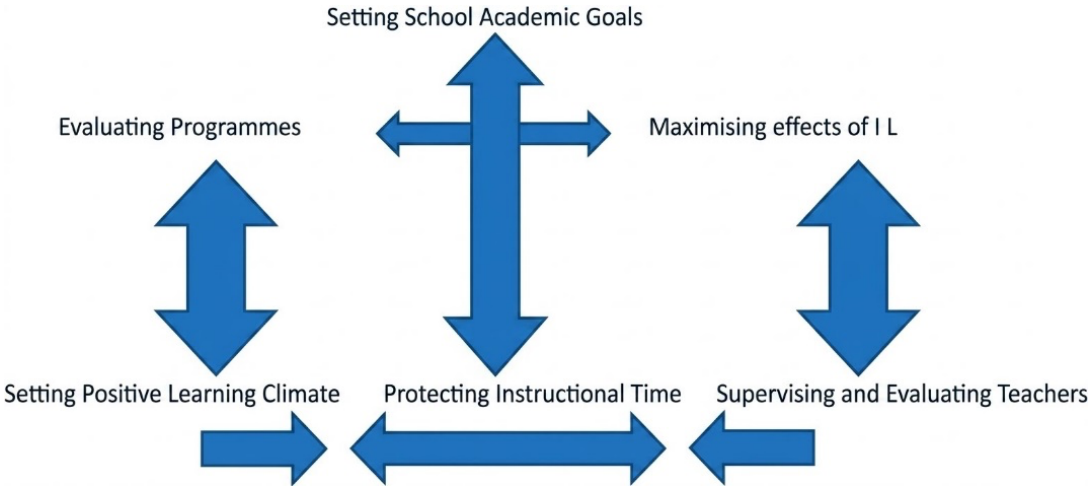
Educational researchers emphasize that success is contingent on support received from districts (Hompashe, 2024; Yingwane, 2025). Fostering mutual respect is essential. Clear role definition and professional independence are fundamental; when district officials intrude into daily operations, hierarchical tension undermines progress (V. M. Ndlovu, 2025). Principals possess understanding of what works best for their institutions (Kgwete, 2024; Shula et al., 2022). Acknowledging their leadership enhances collaboration and improves support (Kgwete, 2024; Maifala, 2024; Shula et al., 2022).

## Theoretical Framework

This article is grounded in instructional leadership, "a model of school leadership in which a principal works alongside teachers to provide support and guidance in establishing best practices in teaching." (Brolund, 2016, p. 42). Principals communicate with staff and set clear goals (Brolund, 2016). SMTs and Subject Advisors provide instructional leadership within their mandates (Bush & Glover, 2009; Grant & Osanloo, 2014). SMTs create conducive environments, set priorities, and ensure delivery (Chabalala & Naidoo, 2021; Klinck et al., 2023; Maifala, 2024; Maja, 2016; Mthiyane, 2019; Ndimande, 2005; Subramoney, 2016; Yingwane, 2025). Subject Advisors focus on supporting teachers through professional development, guidance, and monitoring (DBE, 2013, 2016, 2022b, 2023; Govindasamy & Mestry, 2022).

Instructional leadership is executed through six functions: setting goals, evaluating programmes, maximising effects of instructional leadership, establishing a learning climate, protecting instructional time, and supervising teachers (Hallinger, 2005). Principals and SMTs initiate oversight through coaching, mentoring, and professional development (Mestry, 2017). This study investigated the collective experiences of key stakeholders. Figure 1 illustrates six interconnected functions.

**Figure 1**
*The Interrelationship Amongst Six Functions of Instructional Leadership*




*Source:* Weber's (1996) Model of Instructional Leadership

### Setting School Goals

This function focuses on setting goals. SMTs, led by the principal, provide guidance to shape goals aligning with school philosophy. Subject Advisors collaborate with schools (DBE, 2013). The primary responsibility is to facilitate teaching and learning to improve performance (Krey & Burke, 2005; Robinson et al., 2008).

### Evaluating and Monitoring Programmes

Evaluation and monitoring of programmes is essential for executing the vision. Institutions depend on staff to carry out tasks and implement programmes (He et al., 2024; Krey & Burke, 2005; Weber, 1996).

Instructional leadership includes promotion of learning based on professional inquiry and relationships (Timperley, 2011). Leaders must use data on learner performance to evaluate effectiveness (Hallinger & Murphy, 1992). They manage and evaluate by critically assessing and revising programmes (Chitamba & Jita, 2023). Data analysis enables leaders to set goals, timelines, and strategies (Chitamba & Jita, 2023; Lee et al., 2024).

### Setting Standards for Achievement and Setting Tone and Learning Climate

Educational leaders create a climate for learning where learners set high standards (Weber, 1996). School visits by subject advisors serve as opportunities to inspire teachers (Chitamba & Jita, 2023; Mthembu, 2025; Shushu, 2023). Plans should be underpinned by a clear vision, ensuring alignment with goals (Mbande-Kwayini et al., 2024; Mthembu, 2025).

### Protecting Instructional Time and Programme

Instructional leaders safeguard teaching time and prioritize curriculum delivery (Maponya, 2015). Transformational leadership involves inspiring teachers and learners to achieve goals (Metaferia et al., 2023). Regular communication and visits from district officials enhance self-efficacy (Mbande-Kwayini et al., 2024; Mthembu, 2025). Leadership must promote environments that foster a positive academic climate (Sariakin et al., 2025; Shaked, 2024), cultivating a supportive climate (Hallinger & Murphy, 1992).

### Supervising and Evaluating Teachers

Principals and SMT members supervise and evaluate teachers to maintain order and execution of duties. Personnel selection is important in sustaining instructional programmes, and strategic staffing decisions mitigate challenges (Mestry et al., 2013). SMTs and subject advisors provide staff with opportunities for professional development (Chabalala & Maphalala, 2024; Mestry et al., 2013). District officials provide support whenever principals seek assistance (Mbande–Kwayini et al., 2024; Mthembu, 2025).

### Maximising Effects of Instructional Leadership (IL)

This function underscores the role of instructional leadership programmes and their alignment with goals (Akomodi, 2025; He et al., 2024). Successful implementation depends on effective allocation of human resources and organizational support (Shava & Heystek, 2021). Instructional leaders must possess knowledge of the six core functions: setting goals, evaluating programmes, maximising effects of instructional leadership, establishing a positive learning climate, protecting instructional time, and supervising teachers. This requires leveraging unique strengths of teaching staff to achieve objectives.

## Research Methodology

### Research Paradigm, Approach and Design

The interpretivism paradigm was employed as it seeks to capture participants' first-hand experiences (Creswell & Creswell, 2018; Merriam & Grenier, 2019). Interpretivism is guided by beliefs about the world and how it should be studied (Antwi & Hamza, 2015; Denzin & Lincoln, 2018).

A qualitative research approach was deemed appropriate to engage with SMTs and Subject Advisors. Qualitative research is characterized by interviews, fieldwork, and interaction analysis (Bertram & Christiansen, 2014; Merriam & Grenier, 2019). A case study design was adopted to facilitate data collection (Baškarada, 2014). This approach gains understanding of a phenomenon through multiple collection methods (Coombs, 2022).

### Population and Sampling

Purposive sampling sampled three secondary schools in Dr RSM district. School A is a quintile 3 school from Naledi circuit, School B is a quintile 3 school from Greater Taung Rural circuit, and School C is a quintile 1 school from Molopo Circuit. They were classified as underperforming in Grade 12 for three consecutive years.

Sampling criteria required schools to be operational for fifteen consecutive years with matric results below 65% in two out of three consecutive years. Underperformance was defined as a pass rate of 65% or lower, justified by the Equal Education Law Centre [EELC] (2022). This threshold was confirmed by Bayat et al. (2014), Heystek (2015), Mokoena and Govender (2024), and T. Ndlovu (2025).

In each school, a teacher teaching a matric class and the principal were sampled. Two district officials (Senior Education Specialists) were sampled based on criteria by DBE (2022b, p. 10) regarding monitoring and supporting curriculum implementation.

### Data Collection and Quality Assurance

In-depth interviews were used, and data underwent inductive thematic analysis. Quality assurance criteria of credibility, transferability, confirmability, and dependability were used (Birt et al., 2016). Credibility was confirmed through pre-interviews and prolonged engagements (Elo et al., 2014).

Transferability refers to the extent findings transfer to other settings (Polit & Hungler, 1999). Findings may apply to other districts facing underperformance (Department of Basic Education, 2023). Dependability refers to stability of findings over time, ensured by documenting procedures (Shenton, 2004).

Confirmability allows reliability of results to be scrutinized, minimizing researcher biases (Connelly, 2016). Process transparency improves confirmability (Megheirkouni & Moir, 2023). Interviews were digitally recorded, transcribed, and raw voices presented verbatim.

## Ethical Considerations

The study adhered to ethical considerations to safeguard rights and validity (Gajjar, 2013). Ethical clearance was obtained from Sol Plaatje University ethics committee alongside permission from the North-West Department of Education. Briefing sessions invited voluntary participation (Fleming & Zegwaard, 2018). Informed consent was provided via signed documentation. Pseudonyms safeguarded confidentiality, and data was securely stored.

## Results and Discussion

The research question was: What are the intervention strategies SMTs, and Subject Advisors employ to enhance matric academic performance? The objective was to investigate strategies utilized in Dr. RSM District. Themes emerging from inductive thematic analysis include:

### Psychosocial Support: Motivational Programmes and Initiatives Aimed at Learner Resilience

Most responses identified demotivation as a key factor in underperformance. Participant B from School A stated: "Lack of motivation and frame of reference to want to succeed, so former matriculants who are successful are called to address learners alongside other motivational speakers just to boost them." Participant A from School C shared:

> The learners have lost hope. So, what I'm trying is to bring different faces monthly to come and share a few words with them. I have friends in various professional sectors; I call them to come and motivate my learners.

At School B, a letter invited a university admissions officer to speak to learners.

District SES Participant B highlighted demotivation due to socio-economic challenges. The district collaborates with a local TVET college to invite motivational speakers and states: "Even ourselves as district officials, we visit schools just to encourage learners to focus on their studies." The district leverages school feeding schemes and subsidizes incubation camps with food and transport. District SES Participant A explained:

> We provide socio-psychological assistance to learners to motivate them and show that, even if they lack the means to address their challenges, they can always rely on us as the government for support. Yes, our services are not always available when needed, but we are trying.

### Intensive Pedagogical Support: Academic Camps to Address Content Gaps

Intervention strategies involve winter, autumn, spring, and incubation camps ("academic camps"). District SES Participant A explained that the district organizes autumn, winter, and spring camps and subsidizes year-end incubation camps: "We subsidize the camps and make sure that learners from disadvantaged backgrounds have food to eat when they are there."

Participants highlighted the significance of incubation camps. Participant A from School B described them: "Learners come to school and sleep here. We appoint teachers to supervise and take turns. During this time, a timetable is developed, and learners are drilled in content to prepare them for exams." Schools collaborate to share content. Participant B from School C shared: "We conduct Sunday school alongside weekday extra classes for drilling purposes."

### Institutional Oversight: Human Resources Capacitation and Subject Advisor Support

The district plays an oversight role in managing human resources. District SES Participants A and B highlighted accountability meetings with principals of underperforming schools and ICT integration training for teachers. Strategies ensure necessary materials and resources are available and utilized.

Subject Advisors conduct school visits to monitor teaching and provide support. Participant A from School C explained: “Because of struggles to visit schools, there is one common place where meetings are conducted for principals to share with their teachers.” These initiatives equip teachers and principals with developments to remain aligned with directives from the Department of Basic Education (DBE).

## Conclusion

The research question was: What are the intervention strategies SMTs and Subject Advisors employ to enhance matric academic performance in the Dr. RSM education district? Findings revealed three strategies: motivational programmes and initiatives, academic camps, and human resource capacitation and oversight support.

It is recommended that motivational programmes be extended to remote schools. Subsidies for virtual platforms could address physical access challenges. Inviting district alumni to share experiences may inspire learners. Workshops could empower parents in providing motivational support. Local businesses could contribute through corporate social investment programmes to resource academic camps. Regular training should be offered to teachers and SMTs. Equipping subject advisors with tools such as transportation and technology will enable them to perform roles effectively.

A final recommendation is to encourage research evaluating the effectiveness of these intervention strategies in promoting matric academic performance in Dr RSM education district.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that Grammarly and OpenAI ChatGPT were used solely to assist with background literature exploration, proofreading, and language editing during manuscript preparation. The technology was not used to generate research findings or write core analyses, and the manuscript remains the original work of the authors.

## References

Akomodi, O. A. (2025). The direct and indirect pathways of instructional leadership: Elevating organisational routines and curriculum management in public institutions. *Journal of Educational Leadership and Management, 7*(3), 112–129.

Allais, S. (2014). *Public positions paper: How does matric measure the health of our education system?* Wits Institute for Social and Economic Research. https://wiser.wits.ac.za/system/files/documents/Allais%20-%202014%20-%20Matric%20-%20PublicPositions.pdf

Antwi, S. K., & Hamza, K. (2015). Qualitative and quantitative research paradigms in business research: A philosophical reflection. *International Journal of Research in Business Studies and Management, 2*(4), 1–10.

Bantwini, B. D., & Diko, N. (2011). Factors affecting South African district officials' capacity to provide effective teacher support. *Creative Education, 2*(3), 226–235. https://doi.org/10.4236/ce.2011.23031

Bantwini, B. D., & Moorosi, P. (2018). School district support to schools: Voices and perspectives of school principals in a province in South Africa. *International Journal of Leadership in Education, 21*(6), 757–770. https://doi.org/10.1080/13603124.2017.1394496

Baškarada, S. (2014). Qualitative case study guidelines. *The Qualitative Report, 19*(40), 1–18. https://doi.org/10.46743/2160-3715/2014.1008

Bayat, A., Louw, W., & Rena, R. (2014). Investigating the confluence of factors impacting on underperformance at selected secondary schools in the Western Cape, South Africa. *International Journal of Educational Sciences, 7*(1), 45–55. https://doi.org/10.1080/09751122.2014.11890168

Bertram, C., & Christiansen, I. (2014). *Understanding research: An introduction to reading, writing, and using research in the social sciences*. Juta.

Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. (2016). Member checking: A tool to enhance trustworthiness or merely a nod to validation? *Qualitative Health Research, 26*(13), 1802–1811. https://doi.org/10.1177/1049732316654870

Brolund, L. (2016). Student success through instructional leadership. *BU Journal of Graduate Studies in Education, 8*(2), 41–45.

Bush, T., & Glover, D. (2009). *Managing teaching and learning: A concept paper*. Matthew Goniwe School for Leadership and Governance.

Chabalala, G., & Maphalala, M. C. (2024). Teachers' and middle managers' experiences of principals' instructional leadership towards improving curriculum delivery in schools. *South African Journal of Childhood Education, 14*(1), Article 1483. https://doi.org/10.4102/sajce.v14i1.1483

Chabalala, G., & Naidoo, P. (2021). Teachers' and middle managers' experiences of principals' instructional leadership towards improving curriculum delivery in schools. *South African Journal of Childhood Education, 11*(1), Article 910. https://doi.org/10.4102/sajce.v11i1.910

Chitamba, N., & Jita, L. (2023). Instructional leadership practices: Exploring deputy principal's instructional leadership practices – A Zimbabwean case study. *South African Journal of Education, 43*(4), 1–12. https://doi.org/10.15700/saje.v43n4a2221

Connelly, L. M. (2016). Trustworthiness in qualitative research. *MedSurg Nursing, 25*(6), 435–436.

Coombs, H. (2022). *Case study research defined*. Southern Utah University. https://doi.org/10.5281/zenodo.7604301

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2018). *The SAGE handbook of qualitative research* (5th ed.). SAGE Publications.

Department of Basic Education. (2013). *Policy on the organisation, roles and responsibilities of education districts* (Government Gazette Vol. 574, No. 36324). Government Printers. https://www.education.gov.za/LinkClick.aspx?fileticket=F4jE1wmNQeA%3D

Department of Basic Education. (2016). *Personnel Administration Measures (PAM) (Government Gazette)*. Government Printers.

Department of Basic Education. (2020). *Subject advisor roles and management guidelines*. Government Printers.

Department of Basic Education. (2022a). *Address by the MEC for Education: On the occasion of the release of 2022 NCS results*. North West Provincial Government. http://desd.nwpg.gov.za/wp-content/uploads/2023/01/Matric-Results-Release-Speech-2022-Edited-1.docx.pdf

Department of Basic Education. (2022b). *Final roles and responsibilities of curriculum specialists.* https://www.education.gov.za/Portals/0/Documents/Publications/LRP%20Guidelines/FINAL%20ROLES%20AND%20RESPONSIBILITIES%20OF%20CURRICULUM%20SPECIALISTS.pdf

Department of Basic Education. (2023). *DBE to intensify support to education districts*. https://www.education.gov.za

Ehren, M., Paterson, A., & Baxter, J. (2020). Accountability and trust: Two sides of the same coin? *Journal of Educational Change, 21*(1), 183–213. https://doi.org/10.1007/s10833-019-09352-4

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative content analysis: A focus on trustworthiness. *SAGE Open, 4*(1), 1–10. https://doi.org/10.1177/2158244014522633

Equal Education Law Centre. (2022). *Underperforming schools*. https://eelawcentre.org.za/wp-content/uploads/underperfoming-schools-report-1.pdf

Fleming, J., & Zegwaard, K. E. (2018). Methodologies, methods and ethical considerations for conducting research in work-integrated learning. *International Journal of Work-Integrated Learning, 19*(3), 205–213.

Gajjar, B. (2013). Ethical considerations in research. *Journal of Research in Education and Society, 4*(2), 8–15.

Govindasamy, V., & Mestry, R. (2022). The principal's role in managing curriculum change: Implications for the provision of quality education. *South African Journal of Education, 42*(4), 1–11. https://doi.org/10.15700/saje.v42n4a2294

Grant, C., & Osanloo, A. (2014). Understanding, selecting, and integrating a theoretical framework in dissertation research: Creating the blueprint for your "house". *Administrative Issues Journal: Connecting Education, Practice, and Research, 4*(2), 12–26. https://doi.org/10.5929/2014.4.2.9

Hallinger, P. (2005). Instructional leadership and the school principal: A passing fancy that refuses to fade away. *Leadership and Policy in Schools, 4*(3), 221–239. https://doi.org/10.1080/15700760500244793

Hallinger, P., & Murphy, J. (1992). The principalship in an era of transformation. *Journal of Educational Administration, 30*(3), 77–88. https://doi.org/10.1108/09578239210014341

He, Y., Li, X., & Zhang, J. (2024). Instructional leadership practices and teacher-led growth: Navigating performance monitoring and collaborative school improvement. *Journal of Educational Administration, 62*(2), 145–163. https://doi.org/10.1108/JEA-05-2023-0102

Heystek, J. (2015). Principals' perceptions of the motivation potential of performance agreements in underperforming schools. *South African Journal of Education, 35*(2), 1–10. https://doi.org/10.15700/saje.v35n2a986

Hompashe, D. M. (2024). Does instructional leadership drive educational improvement in South Africa? Evidence from Oaxaca-Blinder decomposition analysis. *Development Southern Africa, 41*(3), 404–426. https://doi.org/10.1080/0376835x.2024.2309455

Kgwete, Y. (2024). Lessons from the ACE school leadership programme on leading during crisis: Enhancing the managerial competencies of South African school principals. *Journal of Educational Studies, 23*(1), 45–62.

Klinck, K., Thutulwa, N., & Pelser, A. M. (2023). Creating a high-performing school management team: Bringing talent to the table for effective service delivery. *Frontiers in Education, 8*, Article 1228181. https://doi.org/10.3389/feduc.2023.1228181

Krey, R. D., & Burke, P. J. (2005). *A design for instructional supervision*. Charles C Thomas Publisher.

Lee, C., Camburn, E. M., & Sebastian, J. (2024). School context, school leaders' data-informed decision making, and student achievement: evidence from Florida. *School Effectiveness and School Improvement, 36*(1), 45–70. https://doi.org/10.1080/09243453.2024.2436889

Maifala, S. (2024). The school management team's role of defining the school vision and mission at a South African rural school. *Perspectives in Education, 42*(2), 223–237. https://doi.org/10.38140/pie.v42i2.7532

Maja, T. S. A. (2016). *School Management Team members' understanding of their duties according to the Personnel Administration Measures* [Master's dissertation, University of Pretoria]. UP Space Institutional Repository.

Manyaka, M. M. (2025). Exploring hybrid leadership's influence on school-based support teams in no-fee primary schools. *South African Journal of Childhood Education, 15*(1), 1–11. https://doi.org/10.4102/sajce.v15i1.1807

Maponya, S. H. (2015). *The role of the principal as instructional leader in improving learner achievement in South African primary schools* [Doctoral dissertation, University of South Africa]. UNISA Institutional Repository.

Mbande-Kwayini, B. G., Rulashe, T., & Ramolobe, K. S. (2024). Education Development Officers as agents for promoting quality education at selected rural schools. *African Journal of Teacher Education and Development, 3*(1), 1–9. https://doi.org/10.4102/ajoted.v3i1.30

Megheirkouni, M., & Moir, J. (2023). Simple but effective criteria: Rethinking excellent qualitative research. *The Qualitative Report, 28*(3), 848–864. https://doi.org/10.46743/2160-3715/2023.5845

Merriam, S. B., & Grenier, R. S. (Eds.). (2019). *Qualitative research in practice: Examples for discussion and analysis* (2nd ed.). Jossey-Bass.

Mestry, R. (2017). Principals' perspectives and experiences of their instructional leadership functions to enhance learner achievement in public schools. *Journal of Education, 69*, 257–280.

Mestry, R., Moonsammy-Koopasammy, I., & Schmidt, M. (2013). The instructional leadership roles of primary school principals. *Education as Change, 17*(1), 49–64. https://doi.org/10.1080/16823206.2014.865990

Metaferia, T., Baraki, Z., & Mebratu, B. (2023). Transformational leadership practices and its influence on teachers' job satisfaction in Addis Ababa government secondary schools. *Cogent Education, 10*(2), Article 2249658. https://doi.org/10.1080/2331186X.2023.2249658

Mkhize, Z., & Otu, M. N. (2018). Understanding black African student attrition in the context of transformation in South African higher education institutions. *Journal of African Foreign Affairs, 5*(1), 149–171. https://doi.org/10.31920/2056-5658/2018/v5n1a8

Mokoena, S., & Govender, S. (2024). Accountable instructional leadership: Assessing the impact of district turnaround plans on underperforming secondary schools. *South African Journal of Education, 44*(2), 112–126. https://doi.org/10.15700/saje.v44n2a2341

Mosibudi, M. P. (2012). Exploring the role of school management teams in enhancing teacher motivation: A South African perspective. *Journal of Educational Administration and Management, 14*(1), 67–84.

Mouton, N., & Malumbete, P. (2023). Exploring the challenges of curriculum advisors in schools in the Vhembe-West district, Limpopo province, South Africa. *South African Journal of Education, 43*(3), 1–9. https://doi.org/10.15700/saje.v43n3a2233

Mpiti, P. T., & Wambu, G. (2023). An investigation into the determinants of underperformance in mathematics among Grade 12 learners in a high school in the Eastern Cape. In D. Ortega-Sánchez (Ed.), *Education annual volume 2023*. IntechOpen. https://doi.org/10.5772/intechopen.109014

Mthembu, P. E. (2025). District leadership strategies and its challenges in supporting teaching and learning: Perspectives of two districts in South Africa. *Frontiers in Education, 10*, Article 150211. https://doi.org/10.3389/feduc.2025.150211

Mthethwa, S. J. (2022). Exploring senior education specialists' roles in developing educators at secondary schools in the Gert Sibande education district. *Independent Journal of Teaching and Learning, 17*(1), 54–68. https://doi.org/10.38140/ijsie.v3i1.2245

Mthiyane, N. C. C. (2019). Context matters: Heads of Department's leadership practices in monitoring and supporting teachers in schools participating in Jika iMfundo. *Journal of Education, 76*, 84–103. https://doi.org/10.17159/2520-9868/i76a05

Munsamy, T. (2023). *The influence of school leaders on teacher performance and learner achievement in public secondary schools* [Master's thesis, University of Johannesburg]. University of Johannesburg Repository. https://ujcontent.uj.ac.za/esploro/outputs/graduate/The-influence-of-school-leaders-on/9955197107691

Munyai, T. R. (2024). *Impact of languages subject advisors' roles on teaching and learning of languages in the intermediate phase, Vhembe District in Limpopo Province* [Master's thesis, University of Venda]. Univen Institutional Repository. http://hdl.handle.net/11602/3342

Ndimande, B. S. (2005). The role of school management teams in enhancing academic performance: A case study of South African schools. *Journal of Educational Leadership and Policy Studies, 8*(2), 123–137.

Ndlovu, T. (2025). Navigating the 65% benchmark: Strategic planning among school management teams in underperforming matric centers. *Perspectives in Education, 43*(1), 89–104. https://doi.org/10.38140/pie.v43i1.7821

Ndlovu, V. M. (2025). *Secondary school departmental heads' experiences of instructional supervision in a South African District* [Master's thesis, University of KwaZulu-Natal]. ResearchSpace@UKZN.

Netshisumbewa, T. (2024). *Impact of teachers in school management team to enhance learners academic performance: A case of Dzondo Circuit, Vhembe District* [Master's thesis, University of Venda]. Univen Institutional Repository.

Ntseto, T. (2015). Teacher motivation and its impact on academic performance: The influence of school leadership practices. *South African Journal of Education, 35*(3), 543–560. https://doi.org/10.15700/saje.v35n3a1078

Polit, D. F., & Hungler, B. P. (1999). *Nursing research: Principles and methods* (6th ed.). Lippincott Williams & Wilkins.

Potterton, M. (2015, December 2). *The primacy of exam results stunts pupils' development*. Mail & Guardian. https://mg.co.za/article/2015-12-02-primacy-of-exam-results-stunts-pupils-development/

Robinson, V. M., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly, 44*(5), 635–674. https://doi.org/10.1177/0013161X08321509

Sariakin, S., Yeni, M., Usman, M. B., Mare, A. S., Munzir, M., & Saleh, M. (2025). Fostering a productive educational environment: the roles of leadership, management practices, and teacher motivation. *Frontiers in Education, 10*, Article 1499064. https://doi.org/10.3389/feduc.2025.1499064

Shaked, H. (2024). Enabling factors of instructional leadership in subject coordinators. *Journal of Educational Administration, 62*(2), 239–254. https://doi.org/10.1108/JEA-09-2023-0223

Shava, G., & Heystek, J. (2021). Managing teaching and learning: Integrating instructional and transformational leadership in South African schools context. *International Journal of Educational Management, 35*(5), 1048–1062. https://doi.org/10.1108/IJEM-06-2020-0294

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*(2), 63–75. https://doi.org/10.3233/EFI-2004-22201

Shula, J., Heystek, J., & Van Wyk, T. (2022). Principals' experiences of servant leadership and accountability for improved school performance in South African schools. *Journal of Educational Leadership and Management, 6*(2), 14–31.

Shushu, H. (2023). The experiences of mathematics subject advisors when conducting school support visits. *Perspectives in Education, 41*(2), 49–61. https://doi.org/10.38140/pie.v41i2.6937

Subramoney, S. (2016). The role of educators in improving academic performance: A case study of content delivery strategies. *Journal of Educational Development, 12*(3), 45–58.

Tabe, H. T. (2023). Teacher quality as a defining factor of student academic performance in schools: The case of a district in South Africa. *African Perspectives of Research in Teaching and Learning, 7*(3), 487–502.

Taylor, N. (2021). The dream of Sisyphus: Mathematics education in South Africa. *South African Journal of Childhood Education, 11*(1), Article 911. https://doi.org/10.4102/sajce.v11i1.911

Timperley, H. (2011). Knowledge and the leadership of learning. *Leadership and Policy in Schools, 10*(2), 145–170. https://doi.org/10.1080/15700763.2011.557519

Weber, J. (1996). Instructional leadership. In S. C. Smith & P. K. Piele (Eds.), *School leadership: Handbook for excellence* (3rd ed., pp. 253–278). Clearinghouse on Educational Management.

West, J., & Meier, C. (2020). Overcrowded classrooms – The Achilles heel of South African education? *South African Journal of Childhood Education, 10*(1), Article 617. https://doi.org/10.4102/sajce.v10i1.617

Yingwane, F. (2025). *The role of school management team in improving the academic performance of Grade 12 learner in Malamulele North-East Circuit, South Africa* [Master's thesis, University of Venda]. Univen Institutional Repository.

# Development of Inclusive Education in Nigeria

Ugonne Ezugwu, University of Regina, Canada

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Efforts to expand equitable access to education in Nigeria have gradually shaped the development of inclusive education policies at both the federal and state levels. Originally rooted in special education approaches, this concept has evolved under the influence of international laws. While previous research has documented these policy efforts, there is limited comparative analysis of how state policies align with federal directives and how contextual factors influence their development. Guided by models of disability, namely the medical, charity, social, and human rights models, this study examines the development of inclusive education policies across six Nigerian states. Using a qualitative document analysis approach, it analyzes federal policy texts and state-level documents selected for their official status, relevance, and recency. The comparative analysis explores various aspects of policy development, including legal frameworks, funding mechanisms, community engagement, and pedagogical innovation. Findings reveal notable differences in the direction of state policies. While national policy advocates a rights-based approach, states implement federal recommendations unevenly, often prioritizing local concerns and resource limitations. Some states, such as Lagos, Edo, Enugu, and Kwara, showcase emerging innovative practices, particularly the integration of digital learning tools to overcome pedagogical barriers. However, implementation remains incomplete, often confined to basic education and rarely extending to private schools. The study indicates that Nigeria's inclusive education policy is gradually moving towards social and human rights models, though achieving full alignment will require clearer federal guidance, improved coordination, and sustained investment to ensure access and participation for all.

*Keywords:* inclusive education, educational policy, Nigeria, special education, educational development

iafor
The International Academic Forum
www.iafor.org

## Introduction

Nigeria appears to place great importance on inclusive education, supported by policy and legislative frameworks at both national and international levels (Federal Ministry of Education, 2023). The concept of inclusive education is embedded in federal and state policies aimed at addressing educational inequalities. Indeed, according to the Federal Government of Nigeria, inclusive education is regarded as a fundamental human right that aligns with international initiatives promoting the right to education for all, such as the United Nations Convention on the Rights of Persons with Disabilities, adopted in 2006 and the 2015 Sustainable Development Goals (FME, 2023).

According to Nigeria's National Policy on Inclusive Education, inclusive education is a concept that involves overcoming barriers to learning opportunities, intending to ensure equitable access to quality education whilst catering for the diversity of learners within the same environment (FME, 2023). These perspectives are based on the need to make education accessible to all children, regardless of their differences, to help them realize their potential. Thus, since the first National Policy on Education (NPE) of 1977, the Nigerian government has affirmed its commitment to guaranteeing access to equitable, high-quality education for all its citizens, regardless of their abilities or disabilities. Section 8 of the policy proposes to provide equal educational opportunities for gifted learners and those with disabilities, so that they may contribute to the nation's socio-economic and technological development (Federal Republic of Nigeria, 1977, Section 8). At the state level, several educational policies have been developed to support marginalized groups, for example, in Plateau, Lagos, Bauchi and Kwara.

Furthermore, this commitment is reflected in numerous federal educational reforms, such as the Universal Primary Education (Federal Republic of Nigeria, 2004; UPE, 1976) programme and the National Policy on Special Needs Education in Nigeria (2015). As the names suggest, these are comprehensive government initiatives aimed at achieving universal basic education and addressing special educational needs.

Recently, a dedicated national policy on inclusive education was adopted in Nigeria and designated the National Policy on Inclusive Education (NPIE) in 2023. This new policy (NPIE) is aligned with global best practice in the field of special education, which emphasizes the education of all children within the same learning environment, with adaptations to take account of each child's individual needs (NPIE, 2023). The policy thus aims to move away from the system of special education for learners with disabilities in favour of an approach that promotes equal access for all to an education free from all forms of discrimination. It also calls on institutions to adapt schools and wider society so that they welcome, manage and celebrate everyone's differences.

However, despite governments' efforts to implement the inclusive education policy across the country, the policy's implementation faces significant challenges.

Firstly, according to numerous studies, the challenge of funding is a fundamental obstacle to the implementation of education policies in the country (Ajuwon, 2008; Eskay & Oboegbulem, 2013; Fabunmi, 2005). According to the authors, the allocation of just 2 percent of consolidated intervention funds to the education of learners with disabilities in basic education does not demonstrate a political commitment to genuine educational inclusion and will not provide the necessary resources. According to Angwaomaodoko

(2023b), most states struggle to access these funds, as the law requires them to contribute 50 percent in the form of taxes or projects before they can benefit from them.

Secondly, there is a gap between policy intentions and their implementation. The inclusive education policy is a federal policy that must be adopted in all states. However, the majority of the states have not yet adopted it. According to an education monitoring report published by UNESCO, only 8 of the 36 states have implemented an inclusive education policy (Pinnock, 2020).

Although the fields of special and inclusive education have been the subject of numerous studies in Nigeria since the first national policy on education in 1977, no study has yet been conducted specifically on the six geopolitical regions. At the regional level, to the best of our knowledge, only one study has analyzed inclusive education practices in South-West Nigeria by exploring the attitudes of teachers and parents towards addressing special educational needs (Adeniyi & Adeyemi, 2015).

Our research is therefore relevant in that it aims to provide an overview of inclusive education in Nigeria, across the six geopolitical zones, to understand its various dimensions within the context of a federal and diverse country. As previously mentioned, the federal states have constitutional authority to oversee and administer education. Through a study of six states (Bauchi, Edo, Enugu, Kaduna, Kwara and Lagos), each situated within a wider geopolitical zone, we seek to understand how inclusive education policies are formulated in the regions, the specific trends and characteristics of these policies, and the factors that underpin them.

In fact, the study will reveal the regional dynamics that influence the development of inclusive education policies.

### Objectives of the Study

- Describe the historical evolution of inclusive education policy in Nigeria.
- Examine the dimensions of inclusive education policy in Nigeria.
- Analyze how inclusive education policies have been developed in selected Nigerian states representing the six geopolitical zones.

### Research Questions

- How is inclusive education developed at the federal and state levels?
- What implementation strategies are adopted by states?

## Literature Review

Research on inclusive education in Nigeria has evolved considerably over the past three decades, reflecting changing national priorities and increasing influence from international frameworks. Early studies primarily focused on the historical foundations of special education and the socio-political factors shaping educational policy, while more recent research has examined policy implementation, stakeholder perceptions, teacher preparedness, legal reforms, and the extent to which inclusive education aligns with global human rights principles. Collectively, the literature demonstrates that although Nigeria has made notable

progress in developing inclusive education policies, significant challenges continue to hinder effective implementation across the country.

The earliest studies emphasized the historical and contextual factors that influenced the development of inclusive education. Obiakor et al. (1990) and Obiakor (1998) argued that cultural beliefs, poverty, political instability, corruption, and regional inequalities constrained the implementation of Nigeria's early special education policies. Similarly, Fabunmi (2005) demonstrated that historical experiences, dating from the colonial period through the post-independence era, continued to shape educational planning and policy development. These studies consistently recommended stronger legislation, increased funding, continuous policy evaluation, and greater recognition of Nigeria's cultural and regional diversity when designing educational reforms.

As inclusive education became increasingly associated with international human rights discourse, researchers shifted their attention toward legal and policy frameworks. Ajuwon (2008) examined the philosophical, sociological, and legal barriers to inclusion and found that inadequate teacher preparation, limited resources, and persistent social prejudice remained major obstacles. Likewise, Akinbola (2010) argued that although international human rights instruments recognize every child's right to education, Nigeria's legal framework has yet to provide sufficient protection for children with disabilities. Other scholars also highlighted structural and socio-political barriers that continue to affect policy implementation, despite repeated revisions of the National Policy on Education and the introduction of initiatives such as the Universal Basic Education programme.

Teacher preparedness has emerged as one of the most widely studied themes in the literature. Several studies indicate that teachers generally support the principles of inclusive education but often lack the professional knowledge and practical skills required to implement it effectively. Surveys conducted by Fakolade et al. (2009), Ajuwon (2012), and Onu et al. (2012) found that professionally trained teachers demonstrate more positive attitudes toward inclusion than untrained teachers and that continuous professional development significantly improves teachers' confidence and competence. However, many teachers still rely on outdated instructional methods and report insufficient training for working with learners with disabilities.

Another recurring concern in the literature is the lack of adequate educational resources and infrastructure. Studies consistently report shortages of specialist teachers, assistive technologies, accessible school buildings, Braille materials, hearing aids, adapted toilets, and other facilities necessary for inclusive learning environments. These shortages undermine policy implementation and contribute to unequal educational opportunities for learners with disabilities. Eskay et al. (2013) further observed that the absence of clear legislative support, combined with insufficient funding and difficulties identifying diverse learning needs, has slowed the transition from segregated special education to genuinely inclusive schooling.

Overall, the literature reveals a gradual shift in Nigerian inclusive education from a predominantly medical and charity-oriented understanding of disability toward social and human rights perspectives. While policy documents increasingly reflect international commitments to inclusion, implementation continues to be shaped by contextual realities, including political commitment, financial capacity, cultural attitudes, institutional readiness, and teacher competence. Most existing studies focus on individual states, specific stakeholder groups, or particular implementation challenges, resulting in a fragmented understanding of

policy development. Comparatively little research has examined how state policies relate to the federal framework or how different contextual conditions influence policy evolution across multiple states. This study addresses that gap by providing a comparative analysis of inclusive education policies across six Nigerian states, thereby contributing to a more comprehensive understanding of the development of inclusive education policy within Nigeria's education system.

## Conceptual Framework

### Models of Disability

The development of inclusive education is linked to four models of disability (Save the Children, 2016). The history of models for understanding disability provides a better understanding of how the concept of inclusive education has gradually Each period has shaped a particular way of viewing learners with disabilities and determining the place they should occupy in schools.

#### *The Charitable Model (1800–1900)*

Religious organizations have exerted a significant influence on the care and education of people with disabilities (Kiran, 2024). Indeed, the charitable model is rooted in the belief that people with disabilities are in a pitiable condition and, consequently, require assistance. From a religious perspective, caring for people with disabilities was seen as a way of putting God's love into practice, sometimes leading to the establishment of special schools or orphanages where basic education could be provided. In some contexts, these services were provided in collaboration with charitable organizations, particularly in the form of alms or temporary aid (Retief et al., 2018). Nevertheless, according to Slee (2011), this practice has been criticized for reinforcing exclusion and discrimination.

#### *The Medical Model of Disability (1900–1970)*

The medical model of disability views disability as a pathology (Chalise, 2024). This model was developed in connection with the education of people with physical and intellectual impairments in residential schools (Nordstrom, 1986), which, according to Parsons (1951) and Slee (2011), primarily required specialized medical intervention for specific impairments. Armstrong et al. (2010) noted that educating children with disabilities was comparable to solving a technical problem that requires identifying the nature of the problem before administering a remedy. The authors argue that, although the approach is rigorous and demanding in terms of human and material resources, it prioritizes the diagnosis of health problems over education. Portalier (2021) noted the role of European specialists in the education of marginalized children within the educational system of the time.

#### *The Social Model of Disability (1980–2000)*

The social model of disability emerged in the 1970s, viewing disability as a barrier arising from the environment (UNESCO, 2020). This model differs from the medical model in that it perceives disability as a mechanism used against minorities to reinforce their discrimination within society.

Consequently, disability has become a social phenomenon, creating barriers for people with disabilities (Kopito, 2024). Thus, social practices act as an obstacle preventing people with disabilities from participating in everyday civic activities (Kiran, 2024). Indeed, the focus is no longer on individual differences and personal deficits, but rather on transforming society so that it meets specific needs; and as long as society does not challenge its view of disability, people with disabilities will continue to face social exclusion (Slee, 2011). The social model is said to have made it possible to challenge the relationship between power, politics and social control, giving rise to the human rights-based approach to inclusion (UNESCO, 2020).

### *The Human Rights Model (2000–2025)*

The human rights approach is primarily enshrined in the United Nations Convention on the Rights of Persons with Disabilities, and is further underpinned by the principle of education that is available, accessible, appropriate and adaptable (UNESCO, 2020). Under the provisions of the Convention, states are responsible for providing an environment conducive to the exercise of the rights of persons with disabilities, including support mechanisms that facilitate their independent living. This model has helped to universalize the rights of these individuals by consolidating them through legal and policy frameworks (Boucher et al., 2011).

In summary, these models have gradually shifted the perspective on disability from a logic of charity and individual correction to an approach based on equality, social justice and the full and equal participation of all learners.

## Methodology

A descriptive qualitative approach was adopted. Qualitative research involves understanding concepts and describing the complex nature of human beings. It seeks to answer questions that are difficult to address using quantitative methodologies (Fortin & Gagnon, 2022). It involves an in-depth examination and explanation of the meaning of a social or human issue through the collection of non-numerical data, from which themes and underlying principles can be deduced. Furthermore, the approach emphasizes the content of the documents in order to reveal themes and groups of salient concepts that help to interpret the situation under study (Fortin & Gagnon, 2022).

### Rationale for the Selection of States

The Federal Republic of Nigeria gained independence on 1 October 1960. Administratively, it comprises thirty-six states and a federal capital territory, divided into six geopolitical zones reflecting diverse ethnocultural affinities (National Population Commission, 2000). According to the National Constitution, each state constitutes a quasi-independent entity with executive, legislative and judicial powers enabling it to formulate and implement public policies, particularly in the areas of education, security, health and the economy (Federal Republic of Nigeria, 1999).

For our study, six states were carefully selected on the basis of the following three criteria: their location within geopolitical zones, the existence of an inclusive education policy, and the adoption of a legislative framework supporting the education of people with disabilities.

Firstly, we selected the states based on their membership of the country's various geopolitical zones. In other words, each of the states represents a geopolitical region of Nigeria, thereby allowing us to take into account the regional dynamics that characterize them. The states include Bauchi in the north-east, Edo in the south-south, Enugu in the south-east, Kwara in the north-central, Kaduna in the north-west and Lagos in the south-west (National Population Commission, 2000).

Secondly, we considered the adoption of the Discrimination Against Persons with Disabilities Prohibition Act 2018, as this is the national legislation that sets out the requirements for the inclusive education of people with disabilities. It lays the foundation for the development of inclusive education policy in the country (Adejoro, 2024). The Act provides for free public education up to secondary level and the provision of infrastructure to support their education, thereby laying the cornerstone for the development of inclusive education policies at the State level. However, of the thirty-six states, only twenty have adopted the Act (Adejoro, 2024).

Finally, our third criterion concerns those states that have developed an inclusive education policy following the adoption of the law (Discrimination Against Persons with Disabilities Prohibition Act 2018). According to a UNESCO Global Monitoring Report on Education, only eight Nigerian states are reported to have drawn up some form of education policy, namely Bauchi, Borno, Edo, Enugu, Jigawa, Kaduna, Kwara and Lagos (Pinnock, 2020).

## Results and Discussion

### Historical Evolution of Inclusive Education in Nigeria

The development of Nigeria's National Policy on Inclusive Education reflects a gradual shift from segregated special education toward a more inclusive education system. The 2017 National Policy on Inclusive Education was developed to harmonize national efforts to improve access to quality education for all learners. It was influenced by international commitments such as the Salamanca Declaration of 1994 and the Dakar World Education Forum of 2000 (FME, 2017). It established inclusive education based on the principles of social justice, equity, non-discrimination, and recognition of learners' diverse educational needs.

As a national commitment, inclusive education is a fundamental philosophy aimed at removing all forms of barriers to education so that every learner can flourish and succeed at school, and the Nigerian government appears committed to ensuring that no citizen is left behind.

Our analysis will be based on five main themes, namely: policy guidelines, teacher training, digital teaching and learning, collaboration, and the issue of policy funding. These themes will enable us to understand the nature of the policies in the six states.

### Policy Orientations

Firstly, the policy orientations underpinning inclusive education policy vary from one state to another. On the one hand, some states have a singular vision of inclusive education and recognize education as a fundamental right of every child. This is the case in Kwara State, where the first policy aimed to include children and young people with albinism in mainstream state schools. The policy provides all the necessary support for the education of

those with albinism, such as the provision of clothing and assistive devices, as well as covering their transport costs. Similarly, the KwaraLEARN initiative aims to transform primary education through the integration of digital tools. Thanks to the programme, around 3,500 teachers in state primary schools have already been trained in the use of digital tools in teaching to improve academic outcomes. Furthermore, some states offer both the full inclusion of pupils with special needs in mainstream schools and segregated education for learners with disabilities. This hybrid approach is implemented in Bauchi, Enugu, Kaduna, Lagos, and Edo. For example, in Bauchi State, educational centres catering exclusively for children and young people with disabilities have been established in every town, such as in Yelwa and Azare. Similarly, in Edo State, there are Edo Progressive Schools for children in rural areas, as well as the EDOBEST initiative, which aims to provide basic education for all through the digitalization of learning.

**Human Resources**

The quality and quantity of teachers are central elements of the policies examined. Several studies have revealed various obstacles related to this issue. According to numerous studies, most teachers lack the skills required to cater to specific needs (Ajuwon, 2008; Angwaomaodoko, 2023a, 2023b). The situation is complex in state schools, where there is a high pupil-to-teacher ratio. For example, the teacher-pupil ratio was 1:73 in Kaduna and 1:79 in Bauchi, respectively (Alabelewe, 2024; Owolabi, 2023).

However, the states do provide initial and in-service training for teachers in special educational needs. In Lagos and Kaduna, teacher training programmes focus on specialization in sign language and the development of positive attitudes towards disability. In contrast, training in Kwara, Edo, and Enugu focuses on the acquisition of technological skills.

**Digital Education**

Digital teaching and learning have been highlighted as essential components of inclusive education for achieving basic numeracy and literacy in certain states. For example, in the states of Kwara, Edo, Kaduna, and Lagos, digital tools are integrated into teaching and learning in State primary schools. In other states, such as Bauchi and Enugu, there are digital schools dedicated to the acquisition of basic numeracy and literacy skills for children aged between six and fifteen. In Enugu State, for example, the construction of 260 Enugu Smart Green Schools in each constituency has already begun. Accessible and free from pre-school through to lower secondary level, the State's objective is to break down inequalities and exclusion in order to achieve basic education based on experiential learning.

**Community Engagement**

Analyses have highlighted the vital role of communities in the state's inclusive policies. These organizations are involved in raising public awareness and organizing training workshops on inclusive education. Their participation has led to successes in the states of Kaduna, Edo, and Bauchi. In contrast to the community-based approach, the states of Kwara and Enugu have relied on the establishment of special committees to promote and manage inclusive education.

Similarly, there have been significant collaborations between the states and private individuals, civil society organizations, international organizations and dedicated associations in the implementation of inclusive education policies across all states. For example, Sightsavers is active in Kaduna, Volontaire International in Enugu, the World Bank and the African Development Bank in Bauchi, the World Bank in Edo, the Albino Foundation in Kwara and Kaduna, and charitable organizations and private individuals in Lagos. These commitments have given rise to numerous initiatives that promote inclusive practices.

**Funding of Inclusive Education**

We found that, except for Bauchi State, which relies solely on international communities, the other states rely partly on federal funds (2% of the federal government's intervention funds) dedicated to special education. However, according to Agwaomaodoko (2023b), these funds are not readily accessible, particularly as the mechanism requires a matching contribution between the federal government and the states in order to benefit from them. As a result, each state appears to be left behind and must therefore go to great lengths to achieve the goal of inclusive education. This is the case in Kwara State, where the government has developed various strategies, including the establishment of a dedicated state fund based on individual or collective donations, the Universal Basic Education Commission fund, and the development of a technical partnership with the private sector.

The funding strategy differs in Lagos State, where, apart from its reliance on federal funds, the Lagos State Government calls on organizations or individuals to adopt schools. According to a recent study, the 2025 budget allocation in Lagos and Edo states falls well short of national and international benchmarks. In other words, 7.8% and 6.4% respectively, compared with the 26% threshold set by UNESCO and Nigeria (Edo State Government, 2025; Nduihe, 2024). However, Enugu funds inclusive education through dedicated state budget allocations. According to the state's annual budgets for 2025, Enugu is among the states that have exceeded the national and international allocation threshold of 26% (Nduihe, 2024). According to data revealed by Nduihe (2024), the state allocated 32.99% of its state budget to education in 2025.

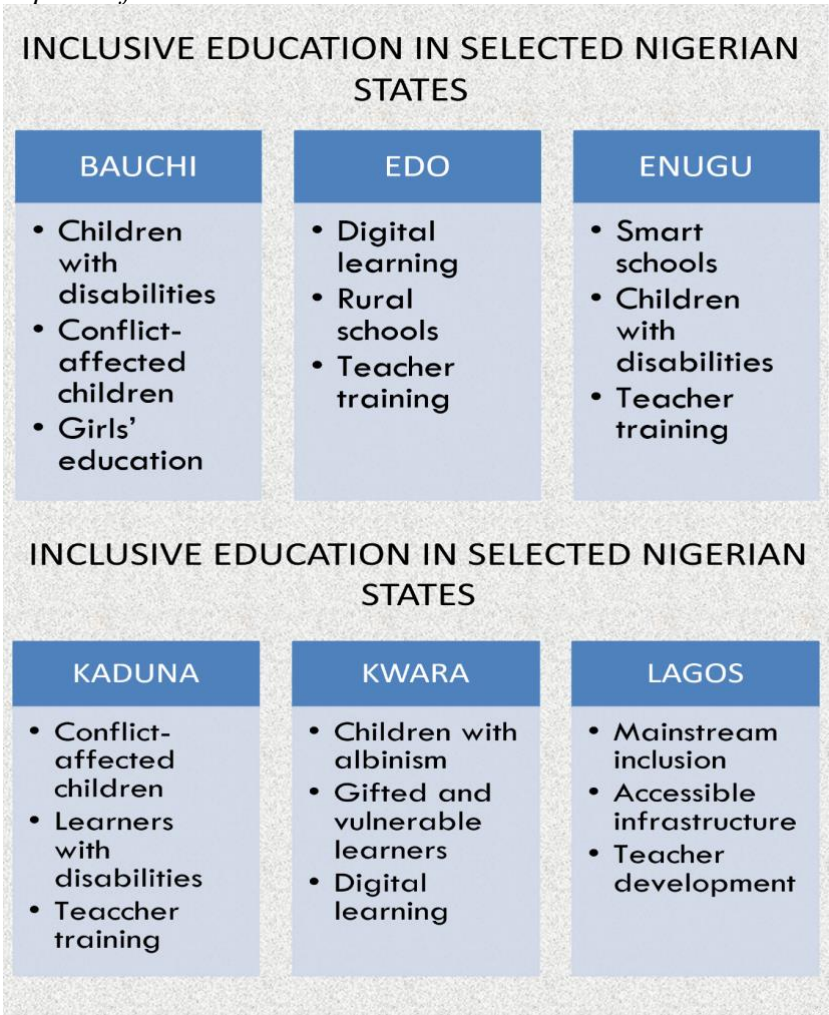
Another new development concerns the approach taken by Kaduna, which has secured a loan at a subsidized rate to promote inclusivity and equity in education (Ministry of Finance, Nigeria, 2025).

**Table 1**
*Historical Development of Inclusive Education Policy in Nigeria*

| Policy Phase | Policy | Year | Policy Orientation |
|---|---|---|---|
| Segregation Phase | NPE | 1977 | Special schools for learners with disabilities |
| Access Phase | UBE | 2000 | Universal access to basic education |
| Integration Phase | SNE | 2015 | Integration of special needs learners |
| Inclusion Phase | NPIE | 2017–2023 | Inclusive education as a human right |

*Source:* Author's adaptation.

**Figure 1**
*Aspects of Inclusive Education in the Selected States*




*Source:* The author.

## Conclusion

This study examined the historical development of inclusive education in Nigeria and compared key dimensions of inclusive education policy at the federal and state levels. The findings show that inclusive education has gradually evolved from a special education approach, which largely separated learners with disabilities from mainstream education, toward a broader vision based on social inclusion and human rights. This transformation has been influenced by international frameworks, including the United Nations Convention on the Rights of Persons with Disabilities, the Salamanca Declaration, and Sustainable Development Goal 4.

The study identified four major dimensions of policy development: legal frameworks, financing, community engagement, and pedagogical innovation. Nigeria's inclusive education policies are supported by national and international legal instruments, but their implementation varies considerably across states. While the federal government provides a broad policy direction, states interpret and apply these recommendations according to their local priorities, resources, and circumstances.

Financing remains one of the major challenges. Although federal funding mechanisms exist, access to these resources is uneven, leaving many states to rely on their own budgets, community support, private contributions, and international initiatives. Some states, particularly Lagos and Edo, have developed innovative approaches to mobilizing resources, while community and international organizations have supported inclusive initiatives in other states. These practices demonstrate opportunities for strengthening and expanding inclusive education financing across Nigeria.

Community participation also plays an important role in supporting inclusion. Parents, individuals, civil society organizations, and local communities contribute to raising awareness, teacher training, and practical support for learners at risk of exclusion. Such involvement is essential because inclusive education cannot be achieved through government policy alone; it requires broader participation and changes in attitudes toward marginalized learners.

Pedagogical innovation, particularly the use of digital technologies, represents another emerging area of progress. States such as Kwara, Enugu, Edo, and Lagos have demonstrated efforts to use digital tools to reduce barriers to learning. However, the effective use of technology requires adequate infrastructure and continuous professional development for teachers and other education personnel.

Despite these developments, the study found that inclusive education policy in Nigeria remains partial. State policies do not consistently incorporate all federal recommendations and are often concentrated on basic education and public schools. The limited evidence concerning private education suggests that inclusive education has yet to become a comprehensive system-wide policy. Consequently, significant differences remain between the national vision of inclusion and its implementation across states.

Overall, Nigeria's inclusive education policy appears to be moving progressively from the medical and charity models toward the social and human rights models of disability. However, achieving meaningful inclusion requires more than the development of policies. The federal government needs to provide clearer guidelines, strengthen coordination among states, improve funding mechanisms, and establish effective monitoring and evaluation systems. States, in turn, need sufficient flexibility and resources to adapt national policies to their local contexts.

The 2023 policy marks a significant departure from the earlier integration and segregation approaches. Rather than expecting learners with special educational needs to fit into an existing system, it promotes an education system capable of adapting to the diverse needs of learners. The policy therefore represents a movement toward the social and human rights models of disability, with the broader goal of ensuring equal access, participation, and quality education for all learners in Nigeria (FME, 2023).

Therefore, for Nigeria to move closer to international standards of inclusive education, future policy efforts should focus on ensuring equal access, meaningful participation, and successful learning outcomes for all learners. A coordinated and sustained national strategy, supported by adequate resources, teacher preparation, community engagement, technological innovation, and systematic monitoring, will be essential for turning the country's commitment to inclusive education into consistent practice.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Grammarly, an AI-assisted writing software, was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. Additionally, Zotero was used to store and manage some of the electronic sources. The author further declares that, apart from Grammarly and Zotero, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from appropriate and systematic conduct of the research.

## References

Adejoro, L. (2024, August 7). *17 states yet to adopt Disability Act – Report*. Punch Newspapers. https://punchng.com/17-states-yet-to-adopt-disability-act-report/

Adeniyi, S. O., & Adeyemi, A. O. (2015). Inclusive Education Practice in Southwestern Nigeria: A Situational Analysis. *Journal of the American Academy of Special Education Professionals*, (page 2).

Ajuwon, P. M. (2008). Inclusive Education for Students with Disabilities in Nigeria: Benefits, Challenges, and Policy Implications. *International Journal of Special Education, 23*(3), 11–16.

Ajuwon, P. M. (2012). Making Inclusive Education Work in Nigeria: Evaluation of Special Educators' Attitudes. *Disability Studies Quarterly, 32*(2). https://doi.org/10.18061/dsq.v32i2.3198

Akinbola, B. R. (2010). The right to inclusive education in Nigeria: Meeting the needs and challenges of children with disabilities. *African Human Rights Law Journal, 10*(2), 457–477.

Alabelewe, A. (2024, February 22). *Concerns over Kaduna's 680,000 out-of-school kids*. The Nation Newspaper. https://thenationonlineng.net/concerns-over-kadunas-680000-out-of-school-kids/

Angwaomaodoko, E. A. (2023a). The Challenges and Opportunities of Inclusive Education in Nigeria. *Traektoriâ Nauki, 9*(7), 1001–1009.

Angwaomaodoko, E. A. (2023b). A Critique of Inclusive Education Policy and the Extent of Its Implementation in Nigeria. *Traektoriâ Nauki, 9*(11), 2001–2008.

Armstrong, A. C., Spandagou, I., & Armstrong, D. (2010). *Inclusive education: International policy & practice*. http://digital.casalini.it/9781446248348

Boucher, N. et al. (2011). Droits humains et personnes ayant des incapacités [Human rights and persons with disabilities]. *Développement Humain, Handicap et Changement Social, 19*(2), 7–18. https://www.erudit.org/en/journals/devhumain/2011-v19-n2 devhumain06809/1086898ar.pdf

Chalise, K. (2024). Models of Disability and Inclusive Education. *International Journal of Disability Research and Policy, 2*(1), Article No. 5.

De Simone, M.E. (2025). *Education reform in Edo State: the EdoBEST program*. World Bank Report. https://documents1.worldbank.org/curated/en/099022825150037517/ pdf/P5058161ea6c4a0261a9be15c55befc82fd.pdf

Eskay, M., & Oboegbulem, A. (2013). Learners with Disabilities in an Inclusive Education Setting in Nigeria: Implications for Administrators. *Online Submission, 3*(5), 313–318. https://eric.ed.gov/?id=ED543446

Fabunmi, M. (2005). Historical analysis of educational policy formulation in Nigeria: Implications for educational planning and policy. *International Journal of African & African American Studies, 4*(2).

Fakolade, O. A., Adeniyi, S. O., & Tella, A. (2009). Attitude of teachers towards the inclusion of special needs children in general education classrooms: the case of teachers in some selected schools in Nigeria. *International Electronic Journal of Elementary Education, 1*(3), 155–169.

Federal Ministry of Education. (2017). *National policy on inclusive education in Nigeria*. https://planipolis.iiep.unesco.org/sites/default/files/ressources/INCLUSIVE-EDUCATION-POLICY.pdf

Federal Ministry of Education. (2023). *National policy on inclusive education in Nigeria*.

Federal Republic of Nigeria. (1999). *Constitution of the Federal Republic of Nigeria 1999. Federal Government Press*. https://nigeriareposit.nln.gov.ng/handle/20.500.14186/466

Federal Republic of Nigeria. (2004). *Compulsory, Free Universal Basic Education Act, 2004*. https://ubec.gov.ng/wp-content/uploads/2021/09/NGA87623-UBE-ACT-2004.pdf

Fortin, M.-F., & Gagnon, J. (2022). *Fondements et étapes du processus de recherche : Méthodes quantitatives et qualitatives* (4e éd.) [Foundations and stages of the research process: Quantitative and qualitative methods] (4e éd.) . Chenelière Education. https://planenigeria.com/wp-content/uploads/2024/11/EXECUTIVE-SUMMARY-IE-POLICY-2023.pdf

Kopito, M. (2024). Inclusive Education: Historical Evolvement and Future Perspective. *Educatia, 21*(28), 38–46.

Michael, E., Onu, V., & Janeth, I. (2012). A Review of Special Education Services Delivery in the United States and Nigeria: Implications for Inclusive Education. *US-China Éducation Review, B9*, 824–831. https://eric.ed.gov/?id=ED537176

National Population Commission. (2000). *Nigeria Demographic and Health Survey 1999*. National Population Commission; United Nations Population Fund; U.S. Agency for International Development. https://dhsprogram.com/pubs/pdf/FR115/FR115.pdf

Nations Unies. (2006). *Convention relative aux droits des personnes handicapées* [Convention on the Rights of Persons with Disabilities]. https://www.ourcommons.ca/Content/Committee/441/FA

Ndujihe, C. (2024). *2025 budget: FG, 22 states allocate N6.13trn to Education*. Vanguard Nigeria. https://www.vanguardngr.com/2024/12/2025-budget-fg-22-states-allocate-n6-13trn-to-education/

Nigeria Ministry of Finance. (2025, August 4). *Nigeria and Kuwait Fund partner to revolutionize education in Kaduna State*. https://finance.gov.ng/nigeria-kuwait-fund-partner-to-revolutionize-education-in-kaduna-state/

Nordstrom, B. H. (1986). *The history of the education of the blind and deaf*. ERIC. https://eric.ed.gov/?id=ED309614

Obiakor, F. E. (1998). Special education reform in Nigeria: prospects and challenges. *European Journal of Special Needs Education, 13*(1), 57–71. https://doi.org/10.1080/0885625980130106

Obiakor, F. E., & Offor, M. T. (2011). Special Éducation Provision in Nigeria: Analyzing Contexts, Problems, and Prospects. *International Journal of Special Education, 26*(1), 25–32.

Obiakor, F. E., Maltby, G. P., & Ihunnah, A. C. (1990, April). *Special education policies in Nigeria: Cultural, socio-economic and political issues* (ERIC Document No. ED320375). Paper presented at the 68th Annual Convention of the Council for Exceptional Children, Toronto, Canada. https://eric.ed.gov/?id=ED320375

Oluremi, F. D. (2015). Inclusive Education Setting in Southwestern Nigeria: Myth or Reality? *Universal Journal of Educational Research, 3*(6), 368–374.

Oluremi, F.D. and Olubukola, O.O. (2013), Impact of facilities on academic performance of students with special needs in mainstream public schools in Southwestern Nigeria. *Journal of Research in Special Educational Needs, 13*, 159–167. https://doi.org/10.1111/j.1471-3802.2011.01228.x

Onu, V. C., Eskay, M. K., Ugwuanyi, L., Igbo, J. N., & Obiyo, N. O. (2012). Professional attributes in teacher preparation for teaching students with hearing impairment: Implications for inclusive education in Nigeria. *US-China Education Review A, 2012*, 3302–3308.

Owolabi, T. (2023, August 9). *BASUBEB debunks report on pupil-teacher ratio*. Daily Post Nigeria. https://dailypost.ng/2023/08/09/basubeb-debunks-report-on-pupils-teacher-ratio/

Pinnock, H. (2020). *Inclusive education in Nigeria: Policy progress weakened by financing*. Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO). https://coilink.org/20.500.12592/3fc18n

Portalier, S. (2021). Trois précurseurs de l'éducation inclusive [Three pioneers of inclusive education]. *La nouvelle revue – Éducation et société inclusives, 89–90*(1), 241–251. https://doi.org/10.3917/nresi.089.0241

Retief, M., & Letšosa, R. (2018). Models of disability: A brief overview. *HTS Teologiese Studies/Theological Studies, 74*(1), Article a4738. https://doi.org/10.4102/hts.v74i1.4738

Save the Children. (2016). *Inclusive education: What, why, and how: A handbook for program implementers*.

Slee, R. (2011). *The irregular school: exclusion, schooling and inclusive education* (1st ed.). Routledge.

Torgbenu, E. L., Oginni, O. S., Opoku, M. P., Nketsia, W., & Agyei-Okyere, E. (2021). Inclusive Education in Nigeria: Exploring Parental Attitude, Knowledge and Perceived Social Norms Influencing Implementation. *International Journal of Inclusive Education, 25*(3), 377–393. https://doi.org/10.1080/13603116.2018.1554715

Ugochukwu, C. O. (2023). *Teachers' Preparedness to Teach in an Inclusive Classroom in Anambra, Nigeria* (Order No. 30987706) [Doctoral dissertation]. ProQuest Dissertations & Theses Global. (2908252257).

UNESCO. (2000). *Le bilan mondial de l'EPT à l'an 2000 au crible* [The Global Assessment of Education under scrutiny] Bulletin publié par l'UNESCO pour le Forum international consultatif sur l'éducation pour tous. https://unesdoc.unesco.org/ark:/48223/pf0000120472_fre

**Contact emails:** ezugwuugonne@gmail.com
ugonne.ezugwu@uregina.ca

# Leading Differently: Experiences of Neurodivergent Education Leaders and Their Professional Journeys

Daniel Beck, University of Chichester, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This study explored the lived experiences and leadership practices of neurodivergent educational leaders within a UK local authority context. Using a participatory exploratory sequential mixed-methods design, it combined survey data, reflexive thematic analysis of semi-structured interviews and co-production of the SUSTAIN reflective framework. Across three phases, the research generated a nuanced account of how neurodivergent leaders experience leadership work, navigate organisational contexts shaped by predominantly neurotypical norms, and sustain their leadership over time. Five interrelated themes were identified: values-led leadership shaped by lived neurodivergent experience; wellbeing-centred leadership shaped by prior harm; cognitive intensity and adaptive scaffolding; the competence paradox; belonging, identity and psychologically safe spaces. Together, these themes informed four interconnected conceptual areas: neurodivergent leadership as a distinctive practice; capability and cost; belonging, identity and organisational context; and conditions for sustainable leadership. The findings informed the co-development of the SUSTAIN Leadership Framework (Supporting Understanding, Sustainability and Thriving in Autistic, ADHD and Identity-aware Neurodivergent Educational Leadership), an applied reflective framework designed to support sustainability, agency and ethical leadership. The study contributes to emerging understandings of neurodivergent leadership by proposing a more relational and ecological understanding of competence, belonging, authenticity and wellbeing.

*Keywords:* neurodiversity, educational leadership, belonging, psychological safety, sustainable leadership

iafor
The International Academic Forum
www.iafor.org

## Introduction

Recent Teacher Tapp survey data reported by Benson (2026) found that 25% of teachers in the UK self-identify as neurodivergent, compared with 17% of educational leaders, raising questions about the experiences and progression of neurodivergent educators within leadership pathways. While research into neurodivergent educators is beginning to emerge, considerably less attention has been given to neurodivergent educational leaders, leaving limited understanding of the factors that support their success, the barriers they encounter and how they experience leadership across their professional journeys (Hewstone, 2024; McKay et al., 2025; Wood et al., 2025).

Although neurodivergent individuals are increasingly visible across education and wider society, research continues to highlight persistent misconceptions, stigma and organisational barriers that can restrict participation and progression (Botha & Gillespie-Lynch, 2022; Kalmanovich-Cohen & Stanton, 2025; Morris et al., 2025; Rosqvist et al., 2020). At the same time, emerging evidence suggests that neurodivergent individuals may bring distinctive strengths to professional practice, including strategic thinking, creativity and problem-solving (Doyle, 2020). However, educational and organisational norms often continue to privilege conventional assumptions about professionalism, communication and leadership, potentially limiting opportunities for neurodivergent leaders to contribute fully and authentically (Brown & Leigh, 2020; Kalmanovich-Cohen & Stanton, 2025; Rosqvist et al., 2020).

These tensions are particularly relevant within educational leadership. Emerging literature suggests that neurodivergent leaders may bring strengths in inclusion, relational practice, innovation and organisational improvement, while simultaneously navigating disclosure concerns, stigma and inconsistent access to support (Hewstone, 2024; Wood et al., 2025). Yet support needs remain poorly understood, and the effort required to sustain successful leadership may remain largely invisible (McKay et al., 2025; The Neurodivergent Coach, 2024).

While competence, belonging and wellbeing have each been associated with positive leadership outcomes and professional flourishing (Allen et al., 2022; Avolio & Gardner, 2005; Maslow, 1968), limited attention has been given to how these experiences interact across individual, relational, organisational and wider cultural systems, or how their interplay shapes the sustainability of neurodivergent leadership. Existing studies have often focused on individual traits or needs rather than examining how competence, belonging and wellbeing emerge through relationships between leaders and their environments (Botha & Gillespie-Lynch, 2022; Brown & Leigh, 2020). Understanding these relationships may help identify the conditions that support authentic and sustainable neurodivergent leadership.

This study explores the experiences and professional journeys of neurodivergent educational leaders in the UK through a participatory exploratory mixed-methods design. It examines how competence, belonging and wellbeing are shaped through interactions between neurodivergence and wider relational, organisational and cultural contexts, while also contributing a co-produced reflective framework designed to support sustainable, values-led leadership practice. In doing so, the study contributes to emerging understandings of leadership as relational, contextual and shaped by diverse ways of thinking, being and leading.

## Literature Review

### Neurodiversity, Leadership and Educational Contexts

The neurodiversity paradigm reframes neurological variation as a dimension of human diversity rather than deficit and recognises lived experience as a legitimate source of knowledge (Rosqvist et al., 2020; Singer, 1998).

Critical Disability Theory (CDT) extends these ideas by challenging assumptions that disability or difference reside solely within individuals, arguing instead that experiences of disability are shaped through social, cultural and institutional structures (Goodley, 2017; Shildrick, 2020). Many difficulties experienced by neurodivergent people may therefore arise through environments designed around neurotypical expectations rather than cognitive difference alone (Goodley, 2017). This perspective is particularly relevant to leadership, where assumptions about professionalism, communication and competence may create both opportunities and barriers for neurodivergent individuals (Brown & Leigh, 2020; Rosqvist et al., 2020).

Both neurodiversity scholarship and CDT question conventional assumptions about competence and professional success, highlighting how such judgements are shaped by social and organisational norms rather than objective measures alone (Goodley, 2017; Rosqvist et al., 2020). Consequently, neurodivergent individuals may be recognised as highly capable whilst simultaneously experiencing significant barriers, unmet support needs or pressure to conceal aspects of their identity (Brown & Leigh, 2020; Kalmanovich-Cohen & Stanton, 2025; Rosqvist et al., 2020).

Researchers increasingly distinguish inclusion from belonging, arguing that participation alone does not guarantee acceptance, authenticity or psychological safety (Allen et al., 2022; Edmondson, 1999). For neurodivergent professionals, belonging may be influenced by disclosure, masking and opportunities for authentic participation (Botha & Gillespie-Lynch, 2022; Brown & Leigh, 2020).

Research examining neurodivergent educators highlights a recurring tension between the strengths they bring and persistent expectations that they adapt to environments shaped around neurotypical norms (Brown & Leigh, 2020; Rosqvist et al., 2020). Educational environments often privilege physical presence, interpersonal immediacy and fixed routines, creating additional demands for neurodivergent staff (Brown & Leigh, 2020). Consequently, neurodivergent educators frequently report fatigue, sensory overwhelm, fluctuating energy levels and the emotional labour associated with masking, while receiving inconsistent organisational support (Brown & Leigh, 2020).

Alongside these challenges, emerging evidence highlights significant strengths. Wood et al. (2025) found that autistic teachers contribute distinctive perspectives, including attention to detail, consistency and meaningful relationships with neurodivergent and marginalised students. Similarly, Guo et al. (2025) found that autistic educators often develop strong rapport through shared understanding and lived experience whilst simultaneously facing pressure to conform to stereotypical expectations of the "ideal teacher."

Many neurodivergent educators report concerns about stigma and professional credibility when considering disclosure (Wood et al., 2025). Research highlights the value of psychological safety, organisational flexibility and supportive mentors in facilitating disclosure and

professional development, although such adaptations are often introduced reactively rather than proactively (Guo et al., 2025). Questions of disclosure, support, belonging and professional credibility may therefore be understood not as separate challenges but as interconnected aspects of negotiating authenticity within educational workplaces.

Educational leadership is widely understood as a dynamic and context-dependent endeavour rather than a static set of innate traits (Day et al., 2016; Leithwood et al., 2006). Contemporary perspectives increasingly emphasise identity, relationships and organisational context in shaping leadership experiences (Day et al., 2020). Bronfenbrenner's (1979) ecological systems theory highlights how leadership experiences are influenced by multiple interacting contexts, including relationships, organisational cultures and wider societal expectations. Maslow's (1968) hierarchy of needs provides a useful lens for considering how safety, belonging and wellbeing may influence leaders' capacity to thrive and sustain their practice over time.

Authentic leadership theory focuses on how leaders understand and express their identities through self-awareness, relational transparency and alignment between personal values and professional practice (Avolio & Gardner, 2005; Walumbwa et al., 2008). However, authenticity may not be equally accessible to all individuals, creating tensions between authentic self-expression and dominant expectations of leadership (Lee, 2021). Similar tensions may affect neurodivergent leaders, who can experience pressure to suppress or mask aspects of their identity to align with neurotypical leadership expectations (Brown & Leigh, 2020; Rosqvist et al., 2020).

Research specifically focused on neurodivergent leadership remains limited. Roese's (2024) Autism Leadership Scale found that autistic participants reported lower leadership self-efficacy and higher levels of social camouflage than non-autistic participants, despite conceptualising leadership similarly. Roberson et al. (2021) further propose that leadership potential emerges through interactions between neurodivergent characteristics and supportive organisational contexts.

Although emerging accounts suggest neurodivergent educational leaders may bring strengths in systems thinking, creativity, inclusion and problem-solving (Hewstone, 2024; Roberson et al., 2021), they continue to navigate masking, disclosure concerns, stigma and the emotional labour of leadership within systems built around neurotypical norms (Rosberg & McNamer, 2025). The literature therefore highlights a recurring tension between the potential value neurodivergent leaders bring and the extent to which prevailing leadership norms enable those strengths to be recognised and sustained. However, important questions remain regarding competence, belonging, wellbeing and authenticity across the professional journeys of neurodivergent educational leaders.

## Methodology

### Research Design

This study adopted a pragmatic mixed-methods approach, informed by an exploratory sequential design (Creswell & Plano Clark, 2018; Morgan, 2007). Pragmatism was selected because it supports methodological flexibility and focuses on understanding phenomena in ways that are useful for addressing real-world problems. This perspective was particularly relevant when researching neurodivergent educational leaders whose experiences are shaped

through interactions between individual identity, relationships and organisational structures, often within systems aligned to neurotypical norms.

Epistemologically, the study adopted a pragmatist fallibilist position in which knowledge is judged by its usefulness and consequences rather than adherence to a single paradigm (Greene, 2007; Kaushik & Walsh, 2019; Morgan, 2007). This position was further informed by the double empathy problem, which reframes communication differences between neurodivergent and non-neurodivergent individuals as relational rather than deficit-based (Milton, 2012; Milton et al., 2022). Knowledge was therefore understood as situated, relational and co-constructed. To support meaningful participation, the study drew upon an adapted version of Lundy's Model of Participation (2007), positioning participants as contributors to interpretation and framework development rather than solely sources of data (Figure 1).

**Figure 1**

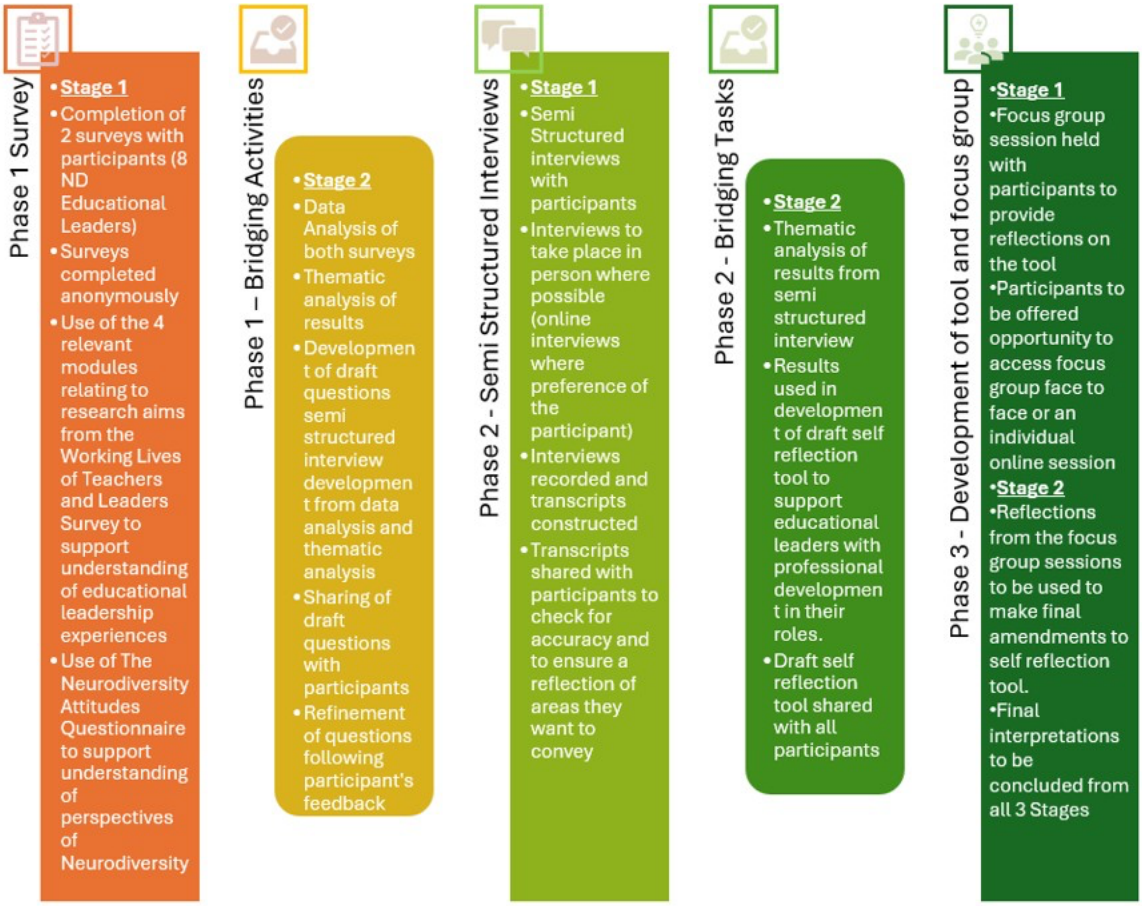
*Adapted Lundy's Model of Participation Checklist to Support an Adult Professional Context*

The study employed an exploratory sequential mixed-methods design across three interconnected phases (Figure 2). Phase 1 comprised survey data collection and analysis. Findings from this phase informed the development of semi-structured interviews undertaken

during Phase 2. Phase 3 involved participant review and refinement of a draft self-reflection tool derived from the findings, introducing a co-productive element that enabled participants to contribute directly to the applied outcome. This design aligned with the study's pragmatic and neurodiversity-affirmative orientation, supporting movement from breadth to depth and ultimately towards practical application (Creswell & Plano Clark, 2018).

**Figure 2**

*Visual Representation of the Three Phases of Exploratory Sequential Design of the Study*




## Participants, Data Collection and Analysis

Eight neurodivergent educational leaders participated in the study (Table 1). Participants represented a range of educational leadership roles and reflected diversity in neurodivergent identity, including formally diagnosed and self-identified neurodivergence. Analysis focused on lived experience rather than diagnostic labels, recognising the heterogeneity of neurodivergence and the limitations of attributing experiences to specific conditions (Milton, 2012; Nind, 2020; Smith & Noble, 2024). The study aimed to generate in-depth, context-specific insight rather than generalisable findings, making a small purposive sample appropriate for qualitative exploration.

**Table 1**
*Overview of Participant Roles in the Study*

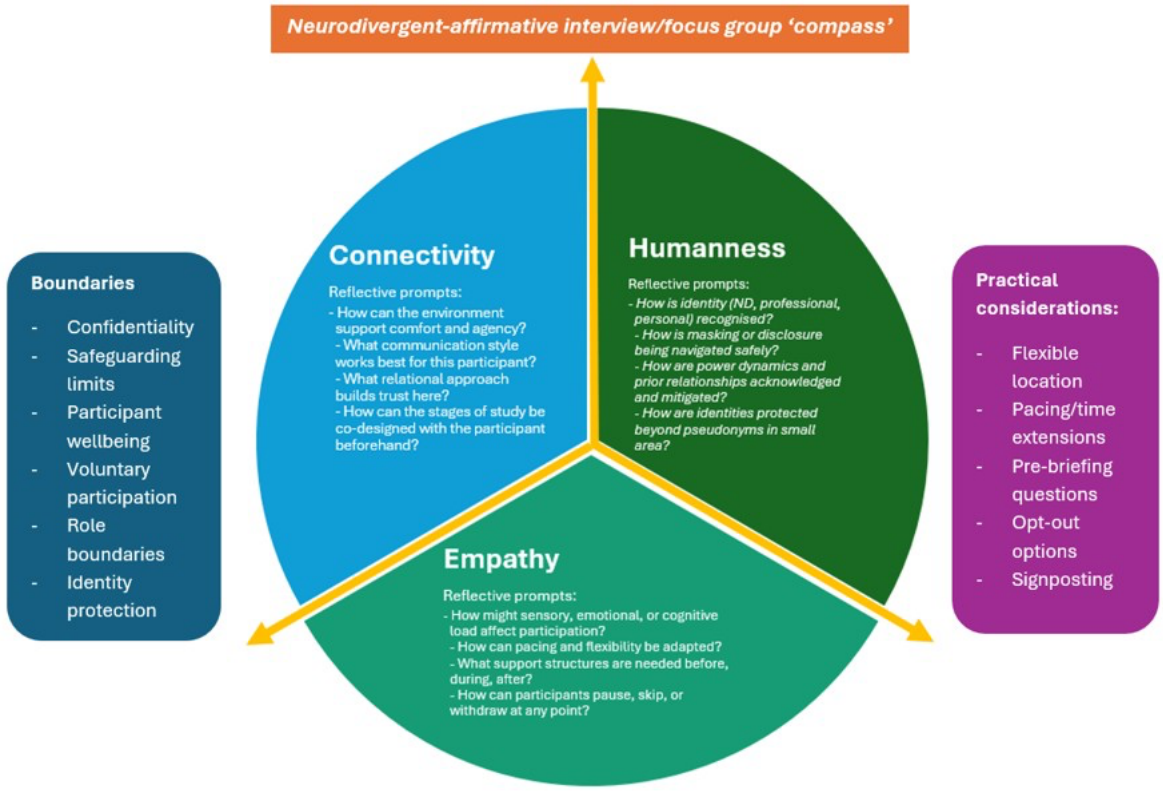
| Educational Leadership Role | Number of Participants |
|---|---|
| Headteacher/Principal | 5 |
| Assistant Headteacher | 2 |
| Special Educational Needs Co-ordinator (SENCo) | 1 |

Participants were recruited through local authority education networks. Surveys explored leaders' experiences, wellbeing and perspectives on neurodivergence. Semi-structured interviews were subsequently analysed using reflexive thematic analysis to identify patterns across participants' experiences. Findings informed development of a self-reflection framework that was reviewed and refined collaboratively with participants during the final phase. This approach enabled participants to contribute not only to data generation but also to interpretation and framework development.

## Ethical Considerations

Ethical considerations were central throughout the study. Ethical approval was granted by the University of Chichester. Approaches to participation, interviewing and framework development were informed by principles of partnership, relational ethics and participant influence. Interviews and focus groups were informed by adapted CHE (Connectivity, Humanness and Empathy) principles that considered communication preferences, sensory environments, masking and disclosure. These principles were operationalised through an adapted reflective compass (Figure 3), guiding decisions regarding engagement, support structures and psychological safety. Ongoing reflexive journaling was used to consider positionality, power dynamics and researcher influence throughout data collection, analysis and framework development. While participant involvement remained bounded by the practical constraints of a small-scale master's study, the research sought to maximise opportunities for participant voice and influence at each stage of the process.

**Figure 3**
*A Neurodivergent-Affirmative Adaptation of Connectivity, Humanness and Empathy (CHE) Principles (Brown & Danaher, 2017) Interview Principles Conceptualised as a Reflective Compass*




## Findings

### Survey Findings

Overall, as visualised in Figure 4, the survey data from both surveys indicated a clear pattern of extended working hours and reported stress existing alongside relatively positive self-reported wellbeing and a sense of control over a high workload, highlighting apparent tensions within participants' experiences. The NDAQ findings indicated strong alignment with neurodiversity-affirming perspectives.

**Figure 4**
*Visual Representation of Key Findings Quantitative Findings From Phase 1 Surveys*

## Semi-structured Interview (SSI) Findings

From the thematic analysis of the SSI, five interconnecting main themes, with supporting subthemes, were identified:

- Values led leadership shaped by lived neurodivergent experience
- Wellbeing centred leadership shaped by prior harm
- Cognitive intensity and adaptive scaffolding
- The competence paradox
- Belonging, identity and psychologically safe spaces

The themes identified through reflexive thematic analysis provide explanatory depth to the survey findings (Figure 5), revealing how participants understood, navigated, and sustained their leadership experiences.

**Figure 5**

*Visual of Integration of Quantitative (Survey) and Qualitative (Semi–structured Interview) Findings*

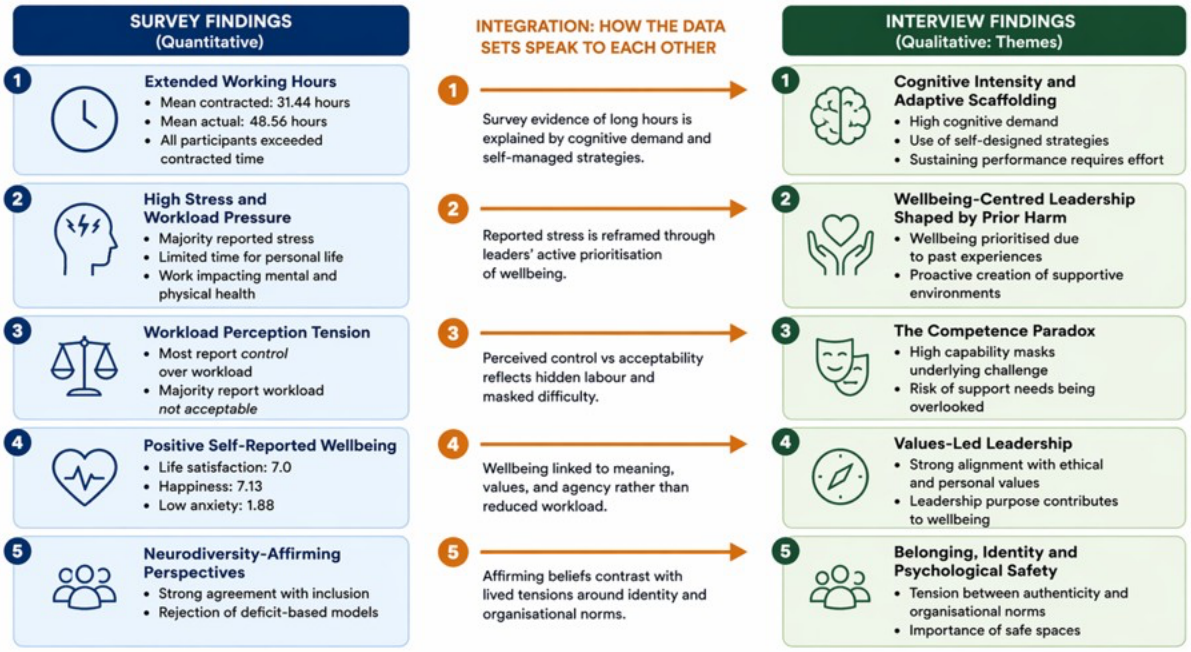



Figure 6 conceptualises the relationships between themes and subthemes identified through reflexive thematic analysis.

**Figure 6**
*Thematic Map Demonstrating the Interconnected Relationships Between Themes and Subthemes in Neurodivergent Educational Leaders' Professional Experiences*

| **Key** | |
|---|---|
| Solid Line | Connection between themes and subtheme |
| Dashed Line | Connection between subtheme |

**Theme: Values Led Leadership Shaped by Lived Neurodivergent Experience**

Participants consistently described leadership as being shaped by lived experience, personal values and a commitment to creating more inclusive environments for others. Leadership decisions were filtered through the lens of pupil impact, with fairness, integrity and compassion repeatedly emphasised. Many connected these values directly to their educational and neurodivergent histories, describing a determination to ensure others did not experience the barriers they had encountered themselves. Participants frequently articulated a desire to prevent harm and "do better for others," positioning lived experience as a source of leadership insight rather than limitation.

These values translated directly into leadership practice. Decision-making was often described as instinctive, relational and values driven. Dani reflected that, "sometimes you've got to stick your neck above the parapet," describing a willingness to challenge systems and advocate for individuals whose needs may otherwise be overlooked. Participants described drawing on their

own sensory, emotional and educational experiences when supporting pupils, families and colleagues. Neurodivergent perspectives were viewed as enhancing rather than constraining leadership. Across interviews, participants emphasised advocacy, authenticity and social justice, suggesting that lived experience was central to how they exercised leadership rather than an aspect of identity separate from it. Leadership therefore emerged not simply as a professional role but as an enactment of deeply held values.

### Theme: Wellbeing Centred Leadership Shaped by Prior Harm

Wellbeing emerged as one of the most interconnected themes within the study. Participants frequently described how previous experiences of unsustainable expectations, burnout and unhealthy workplace cultures had fundamentally shaped their current leadership approaches. Many reflected not only on what had supported them, but on practices they were determined not to reproduce within their own organisations. Dani described previous environments where it was considered "normal" to respond to work communications late into the evening, experiences that informed a conscious commitment to healthier boundaries and more sustainable leadership practices.

Participants consistently framed wellbeing as a leadership responsibility rather than solely an individual concern. Alex described how being "well in myself" was fundamental to leading others effectively, highlighting the reciprocal relationship between personal wellbeing and organisational culture. Boundary setting, flexibility and work-life balance were viewed as essential rather than optional. Across interviews, participants rejected notions of leadership rooted in sacrifice or martyrdom. Sustainable leadership was associated with recognising limits, modelling healthy behaviours and creating conditions in which others could thrive. Rather than presenting resilience as enduring harm, participants described leadership as creating systems that reduce unnecessary harm. Experiences of prior adversity therefore informed more compassionate and sustainable approaches to leadership, wellbeing and workplace culture, demonstrating how lived experience shaped organisational values and leadership decisions.

### Theme: Cognitive Intensity and Adaptive Scaffolding

Participants described cognitive intensity as both a significant leadership strength and a source of challenge. Enhanced pattern recognition, systems thinking and rapid processing were frequently cited as assets within complex educational leadership roles. Participants described being able to "see patterns in data" and simultaneously hold multiple strands of information, enabling strategic thinking and problem solving. However, these strengths were frequently accompanied by cognitive overload, intense rumination and challenges associated with switching off from work-related thinking.

To navigate these demands, participants described developing extensive forms of adaptive scaffolding. These strategies were often highly personalised and developed independently rather than through formal organisational support. Nico explained that "my last hour of the day is always protected," illustrating intentional approaches to managing cognitive load. Participants described using routines, technology, visual systems and structured reflection to externalise demands that could otherwise become overwhelming. AI tools and personalised organisational systems were frequently identified as valuable supports, with one participant describing that "AI has been an absolute game changer." Informal networks were equally important. Leadership meetings were described as feeling "like therapy," providing opportunities for shared understanding and validation. The findings suggest that leadership

effectiveness was not defined by the absence of challenge, but by the extent to which participants could create and maintain scaffolding systems that enabled them to sustain performance whilst managing heightened cognitive and emotional demands.

### Theme: The Competence Paradox

A central finding was the emergence of the competence paradox. Participants frequently described being viewed by others as highly capable, dependable and successful while simultaneously experiencing considerable effort, self-doubt and unmet support needs. Visible competence often resulted in greater responsibility and fewer opportunities for support because colleagues assumed that success equated to an absence of difficulty. Participants were therefore frequently recognised for what they achieved whilst the effort required to achieve it remained largely invisible.

Many participants described experiences associated with perfectionism, impostor syndrome and persistent self-criticism. Nico referred to experiencing “crippling impostor syndrome” despite holding a senior leadership position. Participants reported setting exceptionally high standards for themselves and often becoming their own harshest critics. Luca reflected that “perfectionism has helped me succeed, but it’s also exhausting,” highlighting the tension between achievement and sustainability. Participants rarely sought support, sometimes due to concerns about credibility or fears that acknowledging challenges might undermine perceptions of competence. The findings therefore challenge conventional understandings of capability. Rather than representing a fixed individual attribute, competence appeared relational, contextual and closely linked to available support. Visible leadership success frequently concealed hidden labour, emotional effort and substantial self-management, demonstrating that capability and support needs are not contradictory but can coexist within the same leadership experience.

### Theme: Belonging, Identity and Psychologically Safe Spaces

Belonging was described as a dynamic and relational experience, shaped by organisational culture, trusted relationships and opportunities for authentic participation. Participants viewed identity development as an ongoing process rather than a fixed state. Several explained that leadership had supported greater self-understanding and self-acceptance over time. Nico reflected, “If you'd interviewed me three years ago, I wouldn't have reflected on it this way,” illustrating the evolving relationship between identity, leadership and neurodivergence.

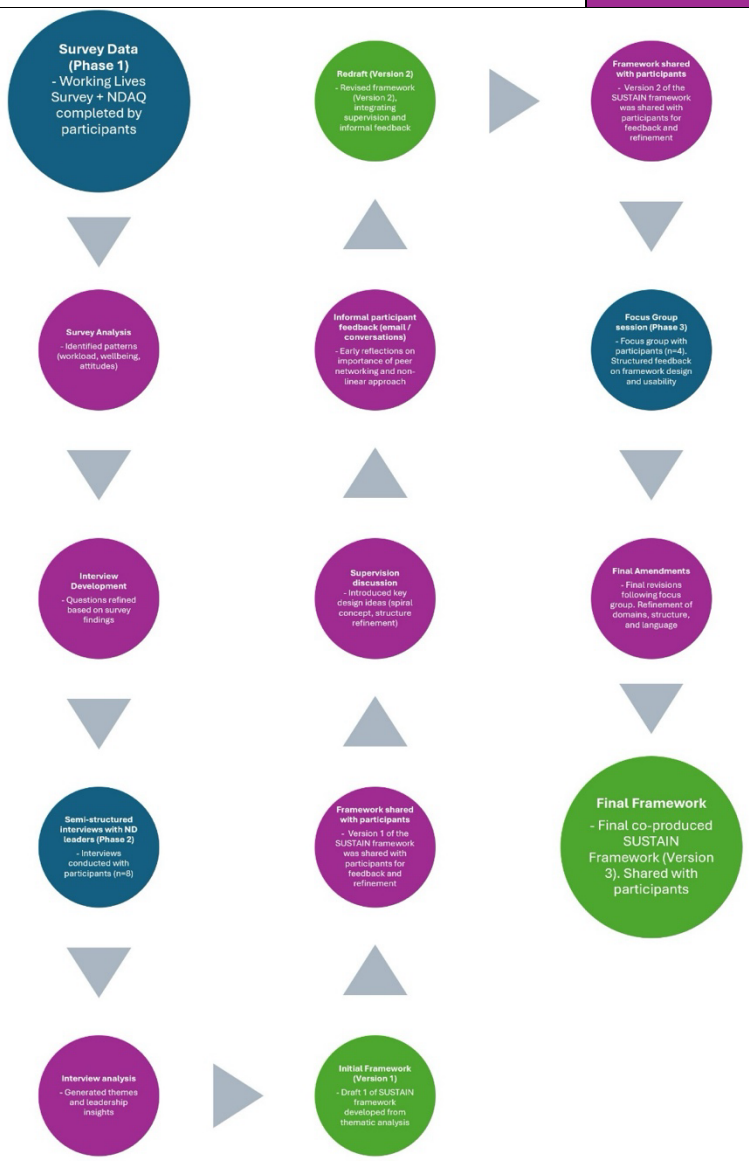
Psychological safety emerged as fundamental to authentic participation. Participants described trusted relationships as enabling openness, vulnerability and professional growth. One participant explained the value of colleagues who could recognise “when I’m about to go into that spiral,” highlighting the importance of relational understanding and anticipatory support. However, participants frequently reported modifying behaviour, masking aspects of identity or making strategic decisions regarding disclosure in less trusted environments. Alex described external leadership spaces as challenging, stating, “In leadership spaces, it’s awful… I constantly feel too much.” Disclosure was therefore rarely straightforward and was often dependent upon trust, context and anticipated consequences. The findings suggest that belonging extends beyond inclusion or representation. Participants emphasised the importance of being accepted, understood and able to contribute without concealing significant aspects of themselves. Belonging, authenticity and psychological safety therefore emerged as closely interconnected conditions influencing both wellbeing and sustainable leadership.

## Framework Development and Focus Group Findings

The development of the SUSTAIN (Supporting Understanding, Sustainability and Thriving in Autistic, ADHD and Identity-aware Neurodivergent Educational Leadership) framework sought to synthesise the findings of this study into a conceptual model for understanding sustainable neurodivergent educational leadership (Figure 7).

**Figure 7**
*Flow Chart Illustrating the Development and Decision-Making Process Underpinning the SUSTAIN Framework*
Key:

| Activity | Colour |
|---|---|
| Phases of Data Collection | |
| Versions of SUSTAIN Framework | |
| Analysis, interpretation, designing | |



Feedback highlighted that leadership experiences were rarely experienced as linear or predictable. In response, the framework adopted a spiral representation (Figure 8), reflecting participants' descriptions of fluctuating confidence, wellbeing, agency and access to opportunities over time. Finally, participants valued the practical and reflective nature of the framework, supporting the inclusion of domain-specific prompts intended to encourage self-reflection rather than assessment or evaluation (Figure 9).

**Figure 8**

*Final Design of the SUSTAIN Leadership Spiral From the SUSTAIN Framework*

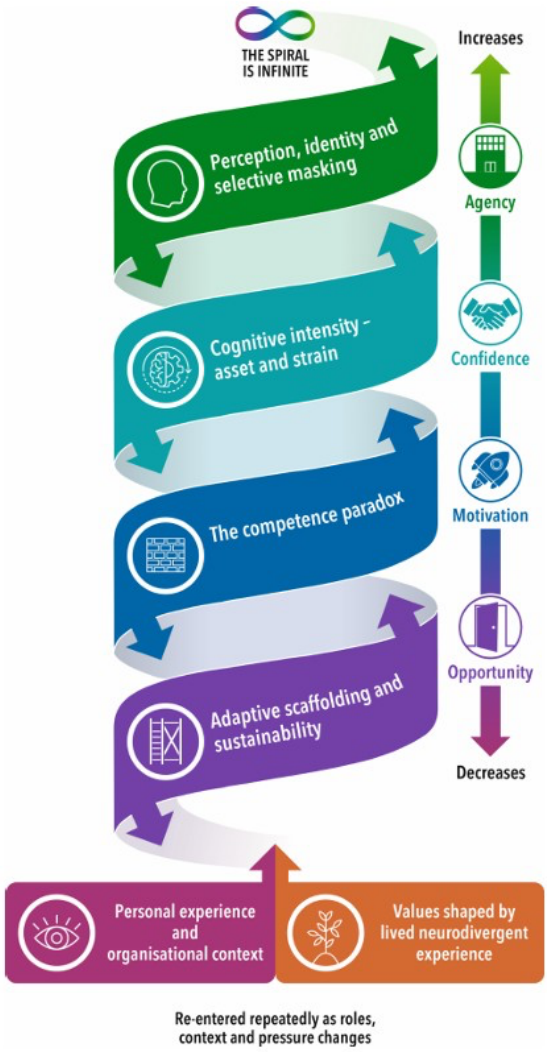


**Figure 9**
*Visual Representation of the SUSTAIN Framework Domains From the SUSTAIN Framework*

## Discussion

### Introduction

The findings suggest that neurodivergent educational leadership cannot be adequately understood through a simple strengths-versus-challenges narrative. Instead, participants described an ongoing interaction between lived experience, leadership practice, wellbeing, identity and organisational context.

The four interconnected findings presented in Figure 10 demonstrate how values-led leadership, capability and cost, belonging and identity, and conditions for sustainable leadership overlap and influence one another. Together, they suggest that neurodivergent leadership is shaped not solely by individual characteristics but through interactions between leaders, organisations and wider social expectations. This aligns with ecological understandings of leadership (Bronfenbrenner, 1979) and supports neurodiversity and Critical Disability Theory perspectives that locate many opportunities and barriers within environments rather than individuals (Goodley, 2017; Shildrick, 2020).

**Figure 10**
*Conceptual Representation of Neurodivergent Leadership Across Four Overlapping Conceptual Areas*

## Neurodivergent Leadership as a Distinctive Practice

Participants consistently described leadership as shaped by lived neurodivergent experience, values and identity. Leadership was frequently connected to experiences of exclusion, misunderstanding or unmet need, resulting in a strong commitment to advocacy and inclusion. Participants repeatedly described doing what was right rather than what was easy, focusing on preventing harm and improving outcomes for others. For example, Indigo reflected that neurodivergence had "empowered me to make sure that no other children are forgotten or missed in the way that I was at school," while Nico described challenging systems "for the one person in the room who that doesn't work for."

These findings extend existing literature surrounding authentic leadership (Avolio & Gardner, 2005; Walumbwa et al., 2008). Rather than authenticity being understood solely as alignment between values and behaviour, participants described leadership itself as being informed by lived experience. Their neurodivergence influenced how they understood barriers, interpreted behaviour and responded to the needs of pupils, families and colleagues. Participants frequently viewed behaviour as communication, emphasised individual circumstances and questioned

assumptions embedded within educational systems. Such findings suggest that neurodivergent leadership may represent a distinctive form of leadership practice shaped by lived experience rather than simply leadership undertaken by neurodivergent individuals.

## Capability and Cost

A second finding concerned the relationship between capability and the hidden effort required to sustain leadership. Participants described significant strengths associated with pattern recognition, systems thinking, empathy and problem solving. However, these strengths frequently co-existed with cognitive intensity, perfectionism and substantial emotional labour. Participants often described developing extensive self-generated scaffolding systems to manage workload, sensory demands and cognitive load.

This finding culminated in what is conceptualised as the competence paradox. Participants were often viewed as highly capable leaders while simultaneously experiencing self-doubt, unmet support needs and considerable personal effort. Nico described experiencing "crippling impostor syndrome," while Luca reflected that "perfectionism has helped me succeed, but it's also exhausting." Visible competence frequently reduced opportunities for support because needs became less visible to others.

These findings challenge assumptions that competence can be understood as a stable and observable individual characteristic. Instead, competence appeared relational and context dependent, influenced by available support, expectations and opportunities for adaptation. This finding echoes neurodiversity scholarship questioning conventional understandings of professional capability and highlights how organisational assumptions may overlook hidden labour and support needs (Brown & Leigh, 2020; Rosqvist et al., 2020).

## Belonging, Identity and Organisational Context

Belonging emerged as a central factor shaping participants' experiences of leadership. Consistent with Allen et al. (2022), participants distinguished between participation and genuine belonging. Holding leadership roles did not necessarily guarantee acceptance, authenticity or psychological safety.

Participants frequently described disclosure as complex, contextual and relational. Decisions to disclose neurodivergence were influenced by trust, organisational culture and anticipated consequences. While some participants experienced environments in which they could participate authentically, others continued to modify behaviour, mask aspects of identity or avoid disclosure. Alex described some leadership spaces as environments where "I constantly feel too much," highlighting continuing tensions between authenticity and dominant expectations of professionalism.

These findings align with Edmondson's (1999) concept of psychological safety and support growing evidence that belonging is not achieved solely through inclusion initiatives. Rather, belonging is shaped by whether individuals can participate without concealing important aspects of themselves. The findings therefore suggest that disclosure may be better understood as an outcome of belonging and psychological safety rather than simply an individual choice. This extends existing neurodivergent workplace research by highlighting the continued importance of identity negotiation even among highly successful educational leaders.

## Conditions for Sustainable Leadership

The findings suggest that sustainable leadership is not simply a product of individual resilience but emerges through the interaction between leader identity, organisational culture, belonging and psychological safety. Consistent with ecological perspectives (Bronfenbrenner, 1979), participants described sustainability as shaped by the environments in which leadership is enacted rather than solely by personal characteristics. Participants frequently rejected leadership models based on sacrifice, overwork and continual availability, instead emphasising sustainability, agency and long-term effectiveness, reflecting principles associated with sustainable leadership (Hargreaves & Fink, 2006).

Many participants described previous experiences of burnout, excessive workload and workplace cultures that normalised unsustainable practices. These experiences informed conscious decisions about boundaries, flexibility and wellbeing. Alex reflected that "I'm well in myself so I can lead them well," while others described the importance of recognising limits, reducing masking and creating conditions in which both staff and leaders could thrive.

The findings also extend authentic leadership theory (Avolio & Gardner, 2005; Walumbwa et al., 2008). Participants suggested that authenticity was dependent upon context and was more readily achieved in environments characterised by belonging and psychological safety (Allen et al., 2022; Edmondson, 1999). These findings informed the development of the SUSTAIN framework and suggest that leadership sustainability is best understood not as an individual responsibility, but as a shared organisational endeavour shaped by the extent to which educational environments recognise and accommodate cognitive diversity.

## Strengths, Limitations and Recommendations

A key strength of this study is its focus on a relatively under-researched area of educational leadership and its use of a participatory, neurodiversity-affirmative methodology. Consistent with calls for greater inclusion of neurodivergent voices within research, participants were positioned as contributors to interpretation and framework development rather than solely subjects of inquiry (Botha, 2021; Nind, 2020). The exploratory sequential mixed-methods design combined survey findings, reflexive thematic analysis and collaborative framework development, enabling both breadth and depth of understanding whilst supporting the co-production of the SUSTAIN framework (Creswell & Plano Clark, 2018; Morgan, 2007). This aligns with neurodiversity research advocating approaches that foreground lived experience and examine environments alongside individuals (Hume, 2024; Karcher et al., 2024).

However, the study does not seek statistical generalisation. The findings represent the experiences of eight self-selecting educational leaders and may not capture the full diversity of neurodivergent leadership experiences across educational contexts. Consistent with qualitative and participatory traditions, the intention was to generate context-specific insight rather than universal claims (Creswell & Poth, 2018; Patton, 2015). As an insider researcher working within a local authority context, positionality may also have influenced participation and interpretation despite the use of reflexive journaling and strategies designed to increase participant influence and reduce power imbalances (Berger, 2015; Mercer, 2007; Nind, 2020).

Future research should explore neurodivergent leadership across different educational phases and organisational contexts, examine the longer-term applicability of the SUSTAIN framework, and investigate how organisational cultures shape belonging, disclosure and sustainability.

Further research may also help extend understanding of the competence paradox and explore how leadership development programmes can recognise lived neurodivergent experience as a source of professional expertise rather than focusing solely on adjustment or accommodation.

## Conclusion

This study explored neurodivergent educational leaders' roles, experiences and wellbeing, while co-developing a reflective tool to support leadership sustainability. The findings suggest that neurodivergence may shape how leadership is enacted, experienced and sustained through distinctive approaches to leadership practice, tensions between capability and cost, experiences of belonging, and conditions for sustainable leadership.

The study demonstrates that neurodivergent educational leaders are not marginal to leadership practice but central to its future development. Leadership emerged as a relational and context-dependent practice shaped by lived experience, values and identity, while strengths associated with neurodivergent cognition, including systems thinking, moral commitment and relational sensitivity, were frequently identified (Bush, 2007; Day et al., 2020; Roberson et al., 2021; Wood et al., 2025).

However, these strengths were not realised independently of context, nor without cost. The competence paradox highlights how visible success can coexist with substantial hidden labour and unmet support needs, suggesting that competence is better understood as relational and context dependent rather than solely an individual characteristic (Brown & Leigh, 2020; Goodley, 2017).

The findings also position belonging, psychological safety and wellbeing as conditions that make sustained leadership participation possible rather than secondary outcomes of effective leadership. In this way, sustainable leadership emerged not from resilience alone, but from the interaction between individual agency and organisational environments that recognise, value and support diverse ways of leading.

The co-produced SUSTAIN framework provides a practical response to these findings. Grounded in neurodiversity-affirmative principles, it conceptualises leadership as dynamic and non-linear, recognising the importance of adaptive scaffolding, reflection and supportive organisational conditions in sustaining leadership over time (Doyle, 2020; Rosqvist et al., 2020).

The disparity between neurodivergent educators and leaders is not simply an issue of representation, but of systemic design. As increasing numbers of neurodivergent educators progress into leadership roles (Benson, 2026), education systems that fail to recognise neurodivergent leadership risk reinforcing exclusion and overlooking perspectives that may contribute to more inclusive and responsive practice. Conversely, systems that legitimise cognitive diversity, adaptive scaffolding and authentic participation may create more equitable and sustainable leadership cultures.

The findings suggest that the future of educational leadership lies not in asking whether neurodivergent individuals can succeed within existing systems, but in understanding how leadership cultures can evolve to recognise, support and benefit from diverse ways of thinking, leading and being.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Microsoft Copilot was used to support proofreading, language refinement, critical reflection and drafting assistance during manuscript preparation. The author reviewed, edited and took full responsibility for all content. The ideas, design, data collection, analysis, findings and conclusions are the author's own.

## References

Allen, K. A., Boyle, C., Sharma, U., Patlamazoglou, L., Pentaris, P., Grové, C., Yared, H., Berger, E., Gamble, N., Morris, Z. A., Finefter-Rosenbluh, I., Morgan, M., & May, F. (2022). Belonging as a core construct at the heart of the inclusion debate, discourse, and practice. In C. Boyle & K-A. Allen (Eds.), *Research for inclusive quality education* (pp. 271–288). Springer. https://doi.org/10.1007/978-981-16-5908-9_21

Avolio, B. J., & Gardner, W. L. (2005). Authentic leadership development: Getting to the root of positive forms of leadership. *The Leadership Quarterly*, *16*(3), 315–338. https://doi.org/10.1016/j.leaqua.2005.03.001

Benson, C. (2026, April 16). *Neurodiversity in the teacher workforce: What new data tells us.* TES Magazine. https://www.tes.com/magazine/analysis/general/neurodiversity-in-teacher-workforce

Berger, R. (2015). Now I see it, now I don't: Researcher's position and reflexivity in qualitative research, *Qualitative Research, 15*(2), 219–234.

Botha, M. (2021). Academic, activist, or advocate? Angry, entangled, and emerging: A critical reflection on autism knowledge production. *Frontiers in Psychology*, *12*, Article 727542. https://doi.org/10.3389/fpsyg.2021.727542

Botha, M., & Gillespie-Lynch, K. (2022). Come as you are: Examining autistic identity development and the neurodiversity movement from an autistic perspective. *Human Development*, *66*(2-3), 93–117. https://doi.org/10.1159/000524123

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Harvard University Press.

Brown, A., & Danaher, P. A. (2017). CHE principles: Facilitating authentic and dialogical semi-structured interviews in educational research. *International Journal of Research & Method in Education*, *42*(1), 76–90. https://doi.org/10.1080/1743727X.2017.137998

Brown, N., & Leigh, J. (2020). *Invisible disabilities: Ableism in education*. British Educational Research Association. https://www.bera.ac.uk/blog/invisible-disabilities-ableism-in-education

Bush, T. (2007). Educational leadership and management: Theory, policy and practice. *South African Journal of Education, 27*(3), 391–406. https://files.eric.ed.gov/fulltext/EJ1066693.pdf

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage.

Day, C., Gu, Q., & Sammons, P. (2016). The impact of leadership on student outcomes: How successful school leaders use transformational and instructional strategies in different contexts. *Educational Administration Quarterly, 52*(2), 221–258. https://doi.org/10.1177/0013161X15616863

Day, C., Sammons, P., & Gorgen, K. (2020). *Successful school leadership.* Education Development Trust. https://files.eric.ed.gov/fulltext/ED614324.pdf

Doyle, N. (2020). Neurodiversity at work: A biopsychosocial model and the impact on working adults. *British Medical Bulletin*, *135*(1), 108–125. https://doi.org/10.1093/bmb/ldaa021

Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350–383. https://doi.org/10.2307/2666999

Goodley, D. (2017). *Disability studies: An interdisciplinary introduction* (2nd ed.). Sage.

Greene, J. C. (2007). *Mixed methods in social inquiry*. Jossey-Bass.

Guo, R., Shen, Y., Hinson-Williams, J., LaPoint, S., McKinnon, R., Wood, R., & Bottema-Beutel, K. (2025). *Navigating teacher agency, empathy, and disability discourses in schools: A reflexive thematic analysis of autistic teachers' perceptions of work experiences in the United States*. OSF Preprints. https://osf.io/j2m38

Hargreaves, A., & Fink, D. (2006). *Sustainable leadership*. Jossey-Bass.

Hewstone, N. (2024). *ADHD heads: How can we utilise neurodiversity in shaping the future of schools?* Diverse Educators. https://www.diverseeducators.co.uk/adhd-heads-how-can-we-utilise-neurodiversity-in-shaping-the-future-of-schools/

Hume, C. (2024). Interpretive interactionism in neurodiversity research. *Qualitative Inquiry*, *30*(1), 45–59.

Kalmanovich-Cohen, H., & Stanton, S. J. (2025). Moving beyond disclosure: Rethinking universal support for neurodivergent employees. *Frontiers in Psychology*, *16*, Article 1547877. https://doi.org/10.3389/fpsyg.2025.1547877

Karcher, K., McCuaig, J., & King-Hill, S. (2024). (Self-)reflection/reflexivity in sensitive qualitative research: A scoping review. *International Journal of Qualitative Methods, 23*, 16094069241261860. https://doi.org/10.1177/16094069241261860

Kaushik, V., & Walsh, C. A. (2019). Pragmatism as a research paradigm and its implications for social work research. *Social Sciences*, *8*(9), 255. https://doi.org/10.3390/socsci8090255

Lee, C. (2021). How do male and female headteachers evaluate their authenticity as school leaders? *Management in Education*, *37*(1), 46–55.

Leithwood, K., Day, C., Sammons, P., Harris, A., & Hopkins, D. (2006). *Successful school leadership: What it is and how it influences pupil learning*. Department for Education and Skills.

Lundy, L. (2007). "Voice" is not enough: Conceptualising Article 12 of the United Nations Convention on the Rights of the Child. *British Educational Research Journal*, *33*(6), 927–942. https://doi.org/10.1080/01411920701657033

Maslow, A. H. (1968). *Toward a psychology of being* (2nd ed.). Van Nostrand.

McKay, A., MacDonald, K., & Longmuir, F. (2025). The emotional intensity of educational leadership: A scoping review. *International Journal of Leadership in Education, 28*(4), 673–695. https://www.researchgate.net/publication/358918571

Mercer, J. (2007). The challenges of insider research in educational institutions: Wielding a double-edged sword and resolving delicate dilemmas. *Oxford Review of Education*, *33*(1), 1–17. https://doi.org/10.1080/03054980601094651

Milton, D. E. M. (2012). On the ontological status of autism: The double empathy problem. *Disability & Society*, *27*(6), 883–887.

Milton, D., Gürbüz, E., & López, B. (2022). The double empathy problem: Ten years on. *Autism, 26*(8), 1901–1903. https://doi.org/10.1177/13623613221129123

Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, *1*(1), 48–76. https://doi.org/10.1177/2345678906292462

Morris, R., Southey, S., & Pilatzke, M. (2025). Autistic-inclusive employment: A qualitative interpretive meta-synthesis. *The British Journal of Social Work,* bcaf024. https://doi.org/10.1093/bjsw/bcaf024

The Neurodivergent Coach. (2024). *Neurodiversity, education leadership, and being human.* https://theneurodivergentcoach.co.uk/education-leadership-and-being-human/

Nind, M. (2020). *Inclusive research: Research methods*. Bloomsbury Academic.

Patton, M. Q. (2015). Qualitative sampling and purposeful sampling. In G. Ritzer (Ed.*), The Blackwell encyclopedia of sociology*. Wiley. https://doi.org/10.1002/9781405165518.wbeoss012.pub2

Roberson, Q., Quigley, N. R., Vickers, K., & Bruck, I. (2021). Reconceptualizing leadership from a neurodiverse perspective. *Group & Organization Management*, *46*(2), 399–423. https://doi.org/10.1177/1059601120987293

Roese, K. (2024). *Measurement of leadership among autistic adults* [Master's thesis, California State University, Fullerton]. https://doi.org/10.5281/zenodo.13401260

Rosberg, D., & McNamer, B. (2025). *Neurodiversity in international school leadership*. The International Educator. https://www.tieonline.com/article/7812/neurodiversity-in-international-school-leadership

Rosqvist, H., Chown, N., & Stenning, A. (2020). *Neurodiversity studies: A new critical paradigm.* Taylor & Francis.

Shildrick, M. (2020). Critical disability studies: Rethinking the conventions for the age of postmodernity. In N. Watson & S. Vehmas (Eds.), *Routledge handbook of disability studies* (2nd ed., pp. 32–44). Routledge.

Singer, J. (1998). *Odd people in: The birth of community amongst people on the autistic spectrum*. University of Technology Sydney.

Smith, J., & Noble, H. (2024). Understanding sources of bias in research. *Evidence-Based Nursing*, *28*(3), 137–142. https://ebn.bmj.com/content/28/3/137.full

Walumbwa, F. O., Avolio, B. J., Gardner, W. L., Wernsing, T. S., & Peterson, S. J. (2008). Authentic leadership: Development and validation of a theory-based measure. *Journal of Management, 34*(1), 89–126. https://doi.org/10.1177/0149206307308913

Wood, R., Gagat-Matula, A., Domagała-Zyśk, E., & Mazur vel Butynski, M. (2025). "We are here and we deserve it": Being an autistic teacher in Poland. *International Journal of Inclusive Education.* Advance online publication. https://doi.org/10.1080/13603116.2025.2518393

**Contact email:** Daniel.Beck@Southampton.gov.uk

# "A Very Good Girl": Experiencing Primary School After Intrafamilial Sexual Abuse From an Adult Male

Sheona Goodyear, University of Birmingham, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

Child sexual abuse is a global concern. Much national and international attention has focused on child sexual exploitation and sexual abuse which is technology assisted or that which happens between children. A high proportion of victims and survivors are girls abused by adult males who are or feel like family. These victims are under-represented in research. Although primary school staff have ideal opportunity to support them, victims rarely disclose sexual abuse experiences at school. I led this research as an insider researcher using heuristic inquiry to provide unique insights into women survivors' experiences of primary school having been sexually abused by an adult male within their family ecosystem. Heuristic inquiry holistically investigates a social phenomenon for which the researcher and research partners share intense personal interest (Moustakas, 2015). Using a best-practice model with a trauma responsive and relational ethic of care, I invited women survivors to share memories and create meaning of what primary school was like for them. Twelve survivor research partners provided rich, nuanced detail of how they experienced school, their protective factors and what might have helped them disclose. Reflexive thematic analysis of the semi-structured interviews and follow-up communications revealed their trauma responses, their efforts to establish autonomy, how they masked what had happened to them and their strong desire to matter in contrast with the sense of not mattering instilled by the sexual abuse and familial responses. The recommendations generated can inform educators, practitioners in other sectors and policy makers to drive enhanced responses for victims.

*Keywords:* intrafamilial child sexual abuse, primary school, heuristic inquiry

iafor
The International Academic Forum
www.iafor.org

## Introduction

This research highlights the real experiences of hidden, ordinary girls sexually abused in their home environment, by family, neighbours and friends. Women survivors of this type of child sexual abuse shared their memories of what it was like being at primary school following intrafamilial child sexual abuse. Although boys and young men are sexually abused, and girls and women can perpetrate abuse, I focused on survivors who were sexually abused as girls aged 12 or younger by men aged 18 or above. In England and Wales, Gekoski, et al. (2016) found this population to be largely ignored in research. In addition to this, evolving Department for Education (DfE) (2025) guidance continues to focus on sexualised behaviour between same aged children, extra-familial and technology assisted sexual abuse. This paper sets out the methodological decisions taken to best meet the aims of the research through heuristic inquiry as described by Moustakas (1990). It explains the ethical practice embedded throughout the research and offers an example of a creatively presented reflection as a way of processing and inciting further research epiphanies. The heuristic analysis of findings is explained and final composite depictions of the survivors' living experiences are shared. This research provides illuminating case studies aimed at teaching and school support staff as well as others involved in child protection, demonstrating likely exemplars of the living experiences of victims while encouraging practitioners to adopt the mindset "it could happen here" (DfE, 2025, p. 17).

## Conceptual Framework

My world view, that the truth about any experience is relative to highly individual circumstances and the sense individuals construct from them, underpinned a constructivist ontology for this research. The aim to interpret individual essential meanings that survivors of intrafamilial child sexual abuse constructed from their memories of attending primary school drove an interpretive epistemology. As an experienced teacher, having taught in primary, further and higher education and managed outstanding special educational needs provision, I have always promoted nuanced, equitable practice. I consider myself to be anti-racist, anti-ableist and anti-sexist and believe that marginalisation is highly linked to invisibility. Although I chose a homogenous population for the research with living experiences similar to mine, it is important to note that among victims of child sexual abuse, people who learn differently and have physical impairments, people from different ethnic backgrounds, and boys are affected by intersecting marginalisation (hooks, 2015) and are more invisible than girl victims. This research was approached from a posthuman perspective which, as Braidotti (2023) described, attempted to challenge colonialism, disrupt marginalisation and create inclusion without privileging any group of people and without privileging humans over other beings or the environment.

To best interpret the essence of living experiences (living signifies that the experience is not time bound—it continues to have impact throughout someone's life) I decided to adopt qualitative research which Denzin and Lincoln had promoted to "make connections between lived experience, social injustices, larger social and cultural structures, and the here and now" (Denzin & Lincoln, 2018, p. 309). I wished to produce new insight to disrupt the silence and silencing still associated with child sexual abuse (Gekoski et al., 2016; Taggart & Wright, 2026), to support prevention of abuse and to facilitate approaches that help child victims to cope with and overcome what has happened to them as soon as is possible.

## Research Background

The United Nations Children's Fund (UNICEF) (2020) and World Health Organisation (WHO) (2017) have both categorised child sexual abuse as a global concern. Prevalence is difficult to establish as it is accepted that many victims only talk about their experiences much later, while some never tell anyone what happened (Jay et al., 2022; Manay & Collin-Vézina, 2021). Globally, it is estimated that 8% of boys and 20% of girls are sexually abused (WHO, 2017). The government in England and Wales have carried out a range of inquiries into differing aspects of child sexual abuse, for example, the Independent Inquiry into Child Sexual Abuse (IICSA) which targeted institutional responses to child sexual abuse. IICSA drew information from Karsna and Kelly's (2021) work which estimated 5% of boys and 15% of girls in England and Wales experience sexual abuse before the age of 16. IICSA called for evidence from survivors, 5,862 of whom consented for their testimony to contribute to the findings. The Children's Commissioner in England had calculated that two thirds of child sexual abuse happened in and around the family home with three quarters of victims being female (Longfield, 2015). Although IICSA was concerned with institutional response to child sexual abuse, half the volunteer participants were survivors of intrafamilial child sexual abuse (Jay et al., 2022) and, underlining the need for research on younger children's experiences, over three quarters of the IICSA survivor participants had been 11 years old or younger when they were first sexually abused.

## Ethical Approach

The University of Birmingham funded this doctoral research and gave ethical approval for the study. The ethical approach taken was embedded throughout all stages of the research, being mindful of unknown survivors in any situation and avoiding harm through trauma responsive practice as advised in the "Survivors" Voices: Survivors Involvement in Research Ladder' (Chevous et al., 2019). Telling a survivor they are safe may make them feel more at risk and watchful (Schachter et al., 2009) so I attempted to make the research feel as unlike being abused as possible. This is in line with the evidence based Handbook on Sensitive Practice for Health Care Practitioners produced by Sneddon, et al. (2016) and other resources produced globally (eg. Quadara and Hunter, 2016; SAMHSA, 2014). I promoted autonomy through the use of transparent information, encouraging personal agency through ongoing consent, openness to questions and allowing time to decide on involvement. Prior to interviews, a relational dialogue built trust and without formally asking, established whether a support network was in place as well as providing signposting to available support. I avoided denying choice and the opportunity to be involved in research as this may also have caused harm. Being open about my experience as a survivor and requesting confidentiality helped develop the reciprocal approach I wished to take which was also enhanced by asking if the survivors wished to choose their own pseudonyms. The aim of the research, to improve understanding for school staff so they are better prepared to support and/or identify victims, resonated with volunteer survivors and in addition to this, being accountable through sharing activities undertaken to disseminate the findings helped to secure ongoing trusting research relationships.

## Methods

In many ways my aspirations for the research aligned with phenomenology. I wished to gain understanding of how survivors interpreted the "meaning, structure, and essence" of their living experiences and to explore their understanding of the phenomenon of intrafamilial child sexual abuse: "how they perceive it, describe it, feel about it, judge it, remember it, make sense of it,

and talk about it with others" (Patton, 2015, p. 115). However, I did not want to interpret the experience to a point where the survivors were excluded from the findings and the essence of the phenomenon related to the experience alone. I also considered my experience an integral part of the research and felt attempting to hide it from participants would be disrespectful. I wanted to be authentic and honest with participants in line with trauma informed practice (Chevous et al., 2019; Quadara & Hunter, 2016; Schachter et al., 2009) and to include my own perspective as a valuable contribution to the research.

## Choosing Heuristic Inquiry

Heuristic inquiry was conceptualised by Moustakas (Moustakas, 1990; Sultan, 2019). It is a qualitative methodology that must be driven and reflexively informed throughout by a researcher's living experience of the phenomenon. It balances looking inwards with looking outwards and explores the essential nature of living experience with ten to fifteen co-researchers or research partners. Unlike in phenomenology, which reduces the essence of experience to a phenomenon and brackets out the individuals who experienced it: the individuals remain an important part of heuristic inquiry findings which generate "shared understanding of the core phenomenon, how it is sensed and experienced, and its significance to oneself, to others, and to the world" (Sultan, 2019, p. 9). I chose to use heuristic inquiry to collect data because it seemed a natural fit for the relational approach to researching with survivors that I wanted to use.

Moustakas (1990) outlined six phases for heuristic inquiry which are: initial engagement, immersion, incubation, illumination, explication and creative synthesis. These are flexible and iterative and can progress in any direction, although the initial engagement, with a socially and globally significant research topic, must originate from personal experience that the researcher feels compelled to investigate. A learning mindset is then required to immersively explore it. Stepping back from immersion for a while, described as incubation, enables subconscious growth and the development of tacit knowledge. While self-reflection and empathic dialogue with research partners enable new insight to emerge in moments of illumination (also known as epiphany). Meanings are explicated into reflexively explored and refined themes and are represented through individual then composite depictions which retain the essence of individuals' experiences. Finally, creative synthesis is required to portray overall interpretations (Moustakas, 1990; Sultan, 2019).

I initially attempted to recruit research partners through a charity gatekeeper and one member of staff from the charity volunteered for the research. While discussing my developing research project in various settings and through engagement with wellbeing and survivor networks, an opportune, information rich sample was created (Patton, 2015). With updated ethical approval, I was able to approach these additional contacts and overall twelve research partners contributed to the research. I invited survivors to engage with in-depth semi-structured interview conversations, through written feedback and with follow up meetings. These conversations were intentionally relational and reciprocal and included self-disclosure with the aim of achieving: congruence, through open, honest, authentic interaction; unconditional positive regard, by valuing and accepting individuals as they are; and empathy by seeking meaning from each other's perspectives (Rogers, 1980).

# Results

## Development of Findings

This paper shows how initial findings, reported at the IAFOR, Paris Conference of Education, 2026 (Goodyear, 2026), were developed beyond key points and themes into composite exemplar portraits of the research partners' experiences. These findings tackle the question: how did the female survivors of intrafamilial child sexual abuse interpret their living experience of primary school attendance?

## Essential Meanings

The research partners reported how they had actively hidden at primary school. They avoided drawing attention to themselves and repeatedly referred to themselves as a "good girl" or "very good girl". Although with suitable conditions, school could provide a safe space or sanctuary, fear dominated the research partners' lives. Their experiences were internalised and manifested as shame leaving them feeling awkward, different, dirty, weird and disgusting. The following three themes were generated from research partners' testimonies:

> How the survivors consciously or subconsciously attempted to assert some *autonomy* in their lives.
>
> How much they were able to identify and present as their *authentic selves.*
>
> How much they felt they *mattered* to the people around them. (Goodyear, 2026)

The research partners' interpretation of their living experience of being at primary school related to experiences of dissociation, hypervigilance and of appeasing people around them. Dissociation is defined by the World Health Organization (WHO) (2022) as a typical result of traumatic experiences, where typical cognitive consciousness can become disrupted or suspended so the individual may lose all awareness of what is going on or feel detached from their body.

Harriet described her experience of being "really dissociated" as: "You know that jelly, you see around tadpoles? … I'm in the middle. And then there's this wall of whatever ~ and then everything else ~ is happening on the other side of the jelly."

Hazel reflected on what being hypervigilant meant for her:

> I was probably hyper vigilant. You know? Where is the next thing going to happen? Because, my amygdala, was still working on the danger. Where's the danger? My, my nervous system was still tuned into, I could get attacked at any moment, you know?

Becky was frustrated with how compliant she had been and remembered "those really horrible, horrible [sentence broken off] and trying to say sorry even though I hadn't done anything!" She also reflected: "I tried to be good. I tried to be good. I tried to do what I was told and go under the radar. And just kind of not be noticeable, really."

## Exemplar Portraits

To finish this report, details of essential meanings from across the research partners have been combined to form three exemplar portraits. These are the final outcome of the heuristic inquiry analysis and present examples constructed as composite characters from the essences of what the research partners said. In reality, these experiences were common across the research partners and any one of them could have presented in these ways at different times.

## The Dissociated Child

The first exemplar is the dissociated child whose coping response to the fear that she is consumed with is dissociation. What this means to her is that she feels numb or not really present, as if she is the dot in the middle of a ball of frogspawn jelly. Most of the time, everything around her happens outside the jelly. She is often confused about what is going on and what she needs to do because she misses conversations, instructions and messages to the extent that she is unaware even of whole school events like sports day. Running is one thing that breaks through the numbness giving her the freedom she craves with a great "whoosh" having been forced to sit still for so long in classrooms. But sports day only deepens her sadness: she does not have the right kit with her, and not knowing why her class are going to the school field, she is still clutching her pencil and rubber. She cannot fly across the grass as she would, showing her swiftness and agility and so even in this arena where she could naturally excel she is denied success.

She finds the classroom interminable. She is the last to "get anything" but the fear of making mistakes and of being noticed means she is desperate to get things right. She has a drive for perfection and cannot wait to achieve her pen license although this is unhelpful when she gets stuck going round in circles, erasing mistakes until her exercise book is full of holes. She is too terrified to be naughty and tries to be "good" while at the same time trying to learn how to be good as she is not really sure what that looks like. She goes to the back of any queue and quietly waits to be noticed. She is not even bothered by bullies because she's so shut down but this dissociated child is lonely as well as frightened. She really wants to be connected with others but is unsure how to achieve this. Sometimes she is able to befriend other more vulnerable children but often she is alone. She feels invisible and has a terrible sense that she does not actually exist. She really wants to "feel loved and cared for" and when school staff are kind to her and try to understand the difficulties she is having, this is a "lifesaving quality".

## The Hypervigilant Child

The second exemplar is a girl whose fear is expressed through hypervigilance. She is not able to trust teachers to keep her safe. Linked to her anxiety and hypervigilance, she avoids skirts and dresses and refuses to brush her hair which is "about not wanting to be pretty". She is occupied by watching for the next danger so, similarly to a dissociated child, she misses instructions and explanations. Because she also worries about drawing attention to herself and is frightened of getting anything wrong, she tries to work out on her own what is required. And because she can read well, she can usually make sense of class work leaving her feeling that her teachers are "a little bit rubbish" when they cannot or do not explain things. Her family are well known at the school but this can also be a barrier to learning, as it is demotivating when she is compared with other family members.

A time when many at the school had a tummy bug has led to her developing extreme anxiety about eating. She avoids foods prepared by other people which makes it difficult for her to engage socially with children outside of school. This also prevents her from going on school visits because she's afraid she will be travel sick.

This hypervigilant child appears self-sufficient but is hiding her feelings and is also very lonely. She does not feel noticed and is certain no one at school would remember her. Watching for danger all the time means that she no longer feels carefree and is rarely seen smiling or laughing. At breaktimes she keeps herself out of the way to avoid attention from bullies and occupies herself by sitting quietly somewhere reading, drawing or making things from leaves and flowers she has collected. She feels she needs to control every aspect of her life to keep herself safe and this is "exhausting". She has been given medication for anxiety but this adds additional risk for her as she is in danger of losing all hope and attempting to harm herself.

### The Appeasing Child

The final exemplar is a girl who feels powerless to deal with complex difficulties she experiences at home but who gains respite from being at school. She tries very hard and is able to achieve at school which gives her a sense of validation. She has a need to be noticed for doing things well and this is expressed by trying to please her teachers, being helpful and friendly. However, she finds it "excruciatingly horrible" if she is seen making a mistake which has led to lack of confidence in maths where she struggles to remember number facts and processes.

This child's anxiety causes regular, excruciating headaches, she is an overachiever because she is working so hard to compensate for her contrasting inner feelings. She feels secretive because she is hiding what has happened at home as well as her sense of shame and disgust. She feels bad and has taken the blame onto herself for what is happening to her. She is desperately afraid of "being found out" and must "just keep coping" somehow.

The appeasing child sometimes has good friends but her relationships are not secure. Sometimes other children physically hurt or ostracise her and she apologises even when she has done nothing wrong in an attempt to appease these bullies. Occasionally, she has become overwhelmed and lashed out and she has also been known to stand up for other vulnerable children and might fight to protect them. She really wants to feel normal and be accepted although inside she feels different, weird and dirty. However, she can escape negative feelings through physical activities like sports or swimming and through arts or creative writing where she can be in the moment and forget everything else.

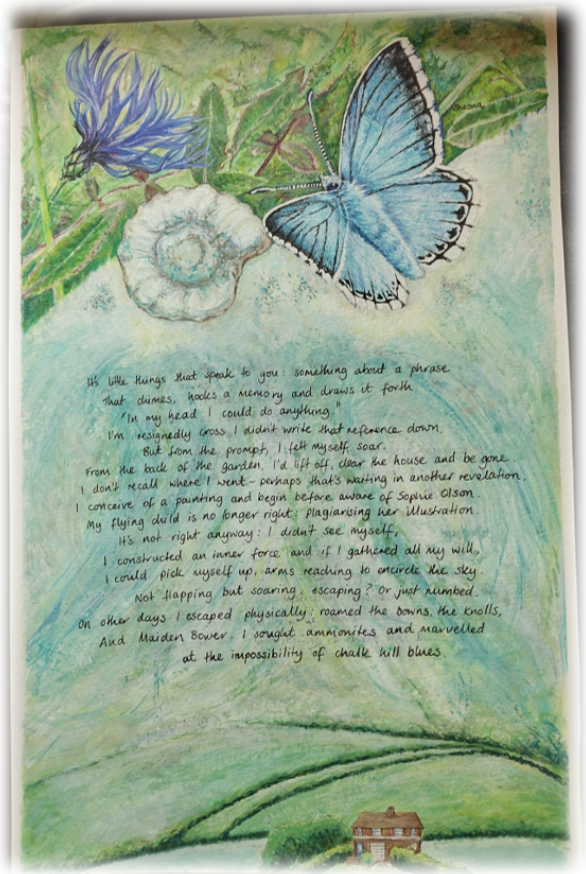
## Conclusion

The choice of heuristic inquiry in this research facilitated rapport between researcher and research partners and helped to avoid power imbalances in the research relationships. I believe the methodology facilitated transformation for me and the research partners. This was enabled through strong connections and trust created through mutual experience; by recognising intersectionality between complex layers of experience, being open and transparent and by accepting the truth of what was shared through mutual belief which avoided mistrust and was affirming and validating. I believe the approach created safe space to be heard and that taking time as the research partners tested how their comments would be received enabled them to share richer details from their lives.

These results were not designed to be generalisable but the transparency and accuracy with which they have been reported should enhance their credibility and trustworthiness. The twelve research partners gave strikingly similar accounts of their experiences of school, demonstrating a shared impact from intrafamilial child sexual abuse which led to them being exceptionally well-behaved and quiet. They had autonomy but they used it to actively hide what had happened to them. When synthesized into exemplar portraits, the experiences correspond with current understanding of reponses to trauma: they presented as children who were dissociated, hypervigilant and appeasing. The conditions constructed within trauma-informed schools which nurture all individuals could act to break down shame and rebuild trust for victims even if they remain hidden. School cultures which value individuals and encourage children to thrive as their authentic selves may reduce the impacts of intrafamilial child sexual abuse and perhaps even create the conditions where victims are enabled to tell someone about it.

These results have been presented as exemplar portraits in the hope that the children portrayed will promote ongoing conversation and thus enhance the research's impact. I would like to finish with an example of illustrated reflective writing consistent with the heuristic inquiry phase of creative synthesis and a final quote to indicate how involving survivors in research can lead to positive growth. Sarah told me that since she began talking about her experience, she is living life in technicolour for the first time, she said: "I've been *existing* for 54 years of my life and now I'm beginning to *live*. In *technicolour*, rather than black and white" (Italics represent Sarah's emphasis).

**Figure 1**
*Epiphany 3: Flying Child, 2025*




*Note:* Number 3 in a series of illustrated pieces of reflective writing representing the experience of dissociation and referencing Katz, et al. (2020) and Olson and Walsh (2024).

## Acknowledgements

I would like to acknowledge the wisdom and generosity of the research partners without whom this research would not have been possible. Also, my supervisors Neil Hall and Sophie King-Hill who have supported me with kindness throughout my doctoral studies.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that no AI or AI-assisted technologies have been used to generate, refine, or correct the content in the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Braidotti, R. (2023). Valedictory lecture: We are rooted but we flow. *Future Humanities, 1*(4), 1–12. https://doi.org/10.1002/fhu2.4

Chevous, J., Perôt, C., & Kennedy, S. (2019). *Survivors Involvement in Research Ladder*. https://survivorsvoices.org/involvement-ladder/

Denzin, N. K., & Lincoln, Y. S. (2018). Introduction: The discipline and practice of qualitative research. In N. K. Denzin and Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research.* 5th Ed. (pp. 1–26) Sage Publications Ltd.

Department for Education (DfE). (2025, September 1). *Keeping children safe in education*. https://assets.publishing.service.gov.uk/media/68add931969253904d155860/Keeping_children_safe_in_education_from_1_September_2025.pdf

Gekoski, A., Davidson, J. C., & Horvath, M. A. H. (2016). The prevalence, nature, and impact of intrafamilial child sexual abuse: findings from a rapid evidence assessment. *Journal of Criminological Research, Policy and Practice*, *2*(4), 231–243. https://doi.org/10.1108/JCRPP-05-2016-0008

Goodyear, S. (2026). Quiet, well-behaved, invisible: How girls experienced primary school after sexual abuse by a man from within their family ecosystem. In *The Paris Conference on Education 2026: Official Conference Proceedings* (pp. 325–338). The International Academic Forum. https://doi.org/10.22492/issn.2758-0962.2026.24

Hooks, B. (2015). Writing without labels. *Appalachian Heritage*, *43*(4), 9–21.

Jay, A., Evans, M., Frank, I., & Sharpling, D. (2022). *The report of the Independent Inquiry into Child Sexual Abuse*. https://www.iicsa.org.uk/reports-recommendations/publications/inquiry/final-report.html

Karsna, K., & Kelly, L. (2021, June). *The scale and nature of child sexual abuse: Review of evidence*. 3rd ed. Centre of expertise on child sexual abuse. https://www.csacentre.org.uk/documents/scale-nature-review-evidence-0621/

Katz, C., Tsur, N., Nicolet, R., Klebanov, B., & Carmel, N. (2020). No way to run or hide: Children's perceptions of their responses during intrafamilial child sexual abuse. *Child Abuse and Neglect, 106*, 104541–104552. https://doi.org/10.1016/j.chiabu.2020.104541

Longfield, A. (2015). *Protecting children from harm: A critical assessment of child sexual abuse in the family network in England and priorities for action*. Children's Commissioner for England. https://www.childrenscommissioner.gov.uk/resource/protecting-children-from-harm/

Manay, N., & Collin-Vézina, D. (2021). Recipients of children's and adolescents' disclosures of childhood sexual abuse: a systematic review. *Child Abuse and Neglect*, *116*(1), 104–192. https://doi.org/10.1016/j.chiabu.2019.104192

Moustakas, C. (1990). *Heuristic research: Design, methodology and applications*. Sage Publications Ltd.

Olson, S., & Walsh, P. (2024). *The flying child: A cautionary fairy-tale for adults*. ZunTold.

Patton, M. Q. (2015). *Qualitative research and evaluation methods*. 4th Ed. Sage Publications Ltd.

Quadara, A., & Hunter, C. (2016, October 26). *Principles of Trauma-informed approaches to child sexual abuse: A discussion paper.* Sydney: Royal Commission into Institutional Responses to Child Sexual Abuse. https://aifs.gov.au/research/commissioned-reports/principles-trauma-informed-approaches-child-sexual-abuse

Schachter, C. L., Stalker, C. A., Teram, E., Lasiuk, G. C., & Danilkewich, A. (2009). *Handbook on sensitive practice for health care practitioner: Lessons from adult survivors of childhood sexual abuse.* 2nd Ed. Public Health Agency of Canada.

Substance Abuse and Mental Health Services Administration (SAMHSA) (2014, July). *SAMHSA's concept of trauma and guidance for a trauma-informed approach*. https://www.nctsn.org/sites/default/files/resources/resource-guide/samhsa_trauma.pdf

Sultan, N. (2019) *Heuristic Inquiry, researching human experience holistically*. Sage Publications Ltd.

Taggart, D., & Wright, K. (2026). Contagions of shame, dignity and connection: Working in the field of childhood sexual abuse. In D. Taggart and J. Stubley (Eds.), *Talking about non-recent child sexual abuse: Survivor, clinician and researcher perspectives.* (pp. 352–366). Routledge.

United Nations Children's Fund (UNICEF). (2020, December). *Action to End Child Sexual Abuse and Exploitation: A Review of the Evidence 2020*. https://www.unicef.org/media/89096/file/CSAE-Report-v2.pdf

World Health Organization (WHO). (2017, October 13). *Responding to children and adolescents who have been sexually abused*. https://www.who.int/publications/i/item/9789241550147

World Health Organization (WHO). (2022, January 1). *ICD-1: International Classification of Diseases 11th Revision*. https://www.who.int/publications/i/item/9789241550147

# Beyond the Classroom: Industry-Mentored Experiential Workshops Transform Interior Design Education

Jeannine Vail, University of North Texas, United States
Kathleen Branning, DLR Group, United States

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Bridging traditional design education with real-world engagement aims to foster deeper human understanding for post-graduate success in the interior design field. This study examines a series of experiential design workshops facilitated by industry professionals. Conducted outside the traditional studio environment, the workshops attracted students from all academic levels, fostering vertical integration and peer learning. Students collaborated through interdisciplinary teams, framed by a hybrid pedagogical model combining Project-Based Learning (PBL) and Experiential Learning, to design a secondary school facility in the United States. With professionals serving as facilitators and mentors, their involvement grounded the project in true-to-life practice, exposing students to realistic, interconnected design influences, including empathy for the client's needs and the importance of communication while achieving solutions through the creative process. Cross-disciplinary collaboration was embedded through the participation of education student majors portraying the "client" and providing user perspectives and goals. The structure reinforced inclusivity through multiple voices and ethical design practices of active listening and accountability to the user. Assessments from surveys, reflective feedback, and observational data revealed a significant pedagogical impact, with 71% of participants reporting that the workshops influenced their preferred learning style. Qualitative findings showed enhanced understanding of end-user needs, appreciation for multiple perspectives gained through collaborative teamwork, and boosted confidence in managing complex, real-world projects. The study suggests that integrating PBL with experiential, industry-led workshops offers an effective model for an enhanced educational experience: expanding learning beyond the studio, strengthening professional connections, and preparing interior design students for career demands in complex, client-centered practice.

*Keywords:* project-based learning, experiential learning, interior design, cross-disciplinary collaboration, industry professionals

iafor
The International Academic Forum
www.iafor.org

## Introduction

The aim of an undergraduate educational process is to develop students into professionals. The integration of collaboration between design professionals and academics helps interior design students achieve an appropriate level of core and generic skills critical to success upon graduation (Rashdan, 2017). One opportunity to ensure effective and valuable training for students is to develop positive partnerships with professionals through extracurricular activities (Roberts & Design, 2007).

The purpose of this research is to examine how educators can bridge traditional design education with real-world industry engagement, with the goal of fostering deeper human understanding and interdisciplinary collaborative learning that prepares students for success after graduation. To implement this goal, a series of five experiential design workshops was developed and facilitated primarily by interior and architecture professionals with support from faculty. Over the course of two semesters, undergraduate students from across all class levels of education collaborated in interdisciplinary teams. The study was framed using a hybrid pedagogical model combining project-based learning with experiential learning, and the vehicle for delivery was a current, tangible design challenge provided by industry experts involving the design of a secondary school facility in the United States. The research aims to explore three questions: (1) How does combining project-based learning (PBL) with experiential learning affect student engagement and comprehension in interior design education? (2) In what ways do industry professionals leading design workshops enhance students' transferable, real-world readiness, particularly empathy and client-centered thinking? and (3) Does relocating design learning activities outside the traditional studio environment influence student motivation, creativity, and engagement with the design process?

## Literature Review

### Experiential Learning in Design Education

Experiential hands-on design studio projects in partnership with industry professionals serve as a pedagogical tool to prepare students for real-world design problems. Literature reveals that learning is maximized when it is self-directed, motivated, and grounded in practical experience that complements formal instruction, particularly when it embraces real-world design (Knowles et al., 2025; Kolb, 1984; Kolb & Kolb, 2005;). Kolb's Experiential Learning Theory (ELT) provides a framework to illustrate how people learn through experience and how this process supports improvement in one's development (Institute for Experiential Learning, 2026). ELT focuses on transforming experience into knowledge through a cycle of four stages: concrete experience, reflective observation, abstract conceptualization, and active experimentation (Kolb, 2015). Studies show that problem-solving involved in hands-on learning activities significantly boosts students' creativity and peer-to-peer learning (Ayob et al., 2012; Buchal, 2011; Vail & Ellis, 2023). A case study with interior design students collaborating with a professional design firm to create a museum found that students enhanced their learning through a better understanding of the design process and professional workflows (Asojo & Vo, 2021).

### Project-Based Learning as a Pedagogical Tool

Project-Based Learning (PBL) is defined by The Buck Institute for Education (2019) as a method of teaching in which students actively engage in real-world projects. Students build

self-awareness of their own learning by exploring multiple perspectives within a collaborative activity (Tamim & Grant, 2013). From a review of the literature, when compared to direct instruction, a PBL approach in teaching has a more positive impact on students' academic achievement (Guo et al., 2020), increasing learned knowledge and skills (Grossman et al., 2019; Wobbe & Stoddard, 2023). The objectives of PBL are to encourage students to collaborate and engage in in-depth inquiry. This approach promotes the development of creativity and critical thinking skills, as well as connections across different subjects. Promoting active participation through PBL aids students in building relationships with the community and increases job placement (Wobbe & Stoddard, 2023). One study with landscape architecture students engaged in a course integrating PBL found students improved their comprehensive ability and increased their motivation to learn when successfully completing project tasks with group interaction (Xi & Wang, 2022). Student perspectives from an additional research study found that, through the PBL process, their motivation for self-directed learning increased. Plus, the teamwork created collaboration and the exchange of knowledge, with insightful feedback from their peers (Tamim & Grant, 2013). The Buck Institute of Education (2019) created a model to aid educators in evaluating and improving PBL practices to help students reach their learning goals, as outlined in Figure 1, through the use of seven essential elements.

**Figure 1**

*Gold Standard PBL – The Seven Essential Project Design Elements*

*Note.* 2019 Buck Institute for Education

## Methodology

### Research Design

The study's design was based on Maxwell's (2008) interactive model of qualitative research, which views design as an ongoing process allowing for the researcher's goals, framework, questions, methods and validity to continually interact, rather than a fixed linear line of steps. This qualitative approach allowed the researchers to explore how educators can use a series of industry-led workshops as a framework to answer their research questions surrounding student engagement, motivation, creativity and comprehension of transferable knowledge and skills from classroom to real-world applications. In keeping with previous studies that sought to explore the benefits of experiential learning with industry design professionals engaged in student instruction, this research combined experiential learning and project-based learning into a single coherent model where experiential learning provided the content and context of students working with real-world professionals, clients and constraints, and project-based learning served as the vehicle for a structured project to drive sustained inquiry and creation. Together, as illustrated in Figure 2, they generate deeper learning by connecting theory to practice, developing skills through authentic problem-solving, increasing engagement by tapping into real-world relevance, and building career readiness through industry-relevant exposure.

**Figure 2**
*Outcomes for Combining Project-Based Learning With Experiential Learning Theory*

### *Participants and Context*

This study was conducted at a public university in the United States, with students enrolled in an interior design program within a college of visual arts and design. Students participated in

a series of five design workshops, focused on educational design. From the current curriculum, the students had previously been exposed to a generic understanding of corporate commercial design but had little knowledge of the education industry segment. The workshops were led by industry professionals, held over the span of one academic year, independent of any course offering. The professionals guided students through the arc of designing a secondary educational facility and the workshops were structured to mirror the actual phases a professional design team moves through on a real project, including sessions on visioning, space planning, project coordination, selection of furniture, fixtures and finishes, plus branded environments. On average, 50 undergraduate students participated in each workshop, with attendance spread fairly evenly across class levels: 29% first-year, 29% second-year, 13% third-year, and 29% fourth-year students. Leading the workshops, the project partnered with 18 design professionals through the Association for Learning Environments, which is an interdisciplinary design association of architects, interior designers, and engineers. Furthermore, 10 manufacturing and multi-line representatives from the furniture, fixture and finishes industries joined the workshops, bringing physical components to instruct students on the impact of this added layer of reality into the design evolution of an environment. An additional ten students from the university's education, art education, design communication and user experience/user interface programs attended the workshops to serve in the client role for the project and lend their vision, needs and expertise to the design discussion. During the workshops, the researchers recorded observations, and at the conclusion of the series, a survey questionnaire and one-on-one interviews were conducted to capture student participants' learning experiences to address the research questions.

## Project Development and Implementation

Instead of students designing for an abstract user, the project focused on designing a facility for a real high school located 50 kilometers from the university and recently destroyed by a tornado. This school's demographics include a 99% minority enrollment with 81% of students economically disadvantaged, and 66% of the population at risk of dropping out. Students were tasked with creating designs for a real, vulnerable community that was healing from trauma, requiring students to design with empathy and encouraging them to better understand the experiences of the people who would use the building.

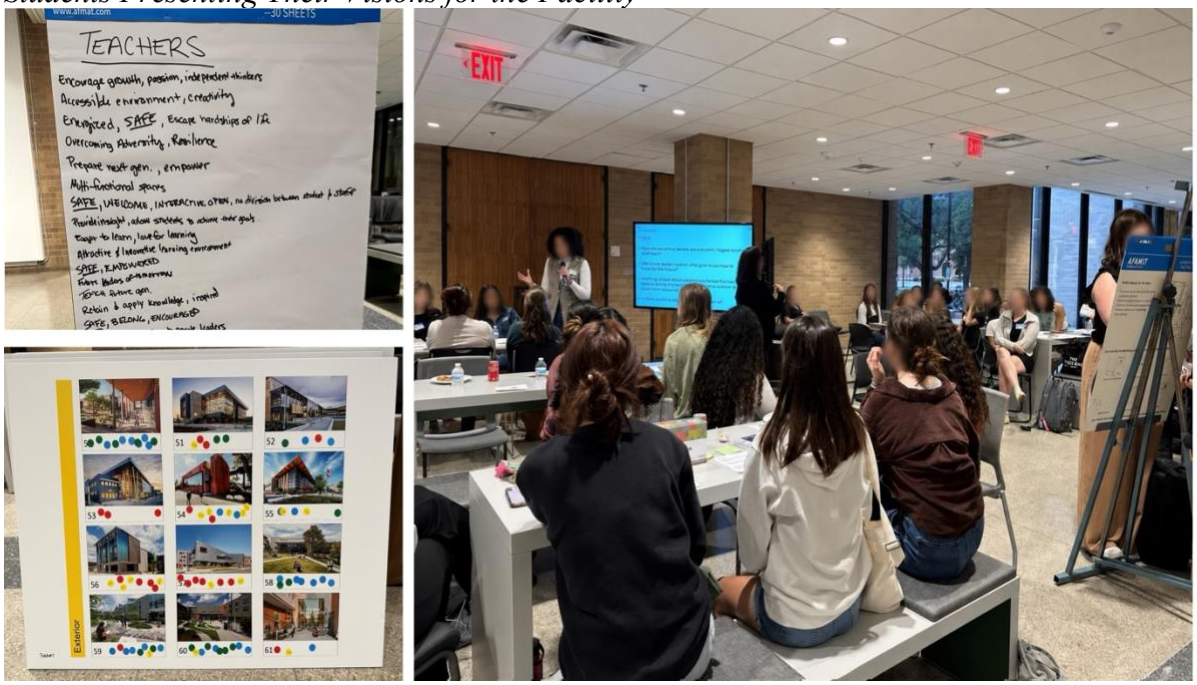
### *Workshop One*

The first workshop focused on Visioning and laid the groundwork for the workshop series. Students were introduced to their "client" and given the key factors about the high school and the reason for necessitating the renovation: the tornado, but also demographics, past history of the school, and their core values. Students were split into groups that they would stay with for the full workshop series. Comprised of varied grade levels and other majors, portraying the "client," to ensure they were gaining real-to-life experience as much as possible. In firms, design teams are made up of a vast range of experience, disciplines, and personalities; thus, this workshop was structured to best replicate that experience. The groups were tasked with answering questions as either an educator, staff member, or student, and placing themselves in the shoes of the clients described, while also using their foundational knowledge of being students and potentially future educators and designers. They were asked to identify goals and foundational needs, as well as envision wishes and dreams for the school and each space within. As pictured in Figure 3, students were additionally led through a red dot/green dot exercise that is often conducted with clients to gauge their design style and see what they gravitate towards or away from. They were given images of exteriors of schools, libraries, classrooms, and

cafeterias. For each category, students placed a red dot on images they did not like and a green dot on images they loved. This culminated in group sharing of design drivers and vision statements that would carry through to the following workshops and guide their designs.

**Figure 3**

*Students Presenting Their Visions for the Facility*




***Workshop Two***

Space Planning was the focus of workshop two. The students were reminded of their vision statement and then, at random, given a classroom wing of the school with a specific area of focus: Career and Technology Education, Science, Art, or Special Needs. The goal was to encourage students to think outside of their existing knowledge, while showing that despite using the same basic elements - for example, classrooms, restrooms, and teacher workrooms - every solution would look different but also potentially shine light on commonalities of their space planning and discuss convergence points in their decision-making. While every group was required to use the same core planning blocks, at realistic square footage, they were also given design freedom in allowing for additional "specialty" programming that they determined their user might need or want based on their vision. Figure 4 shows students working in teams to develop their design for their given space. Education majors in each group grounded the ideas, and interior design majors pushed boundaries and used their prior space planning experience to make a case for each decision-making component. Just as in the prior session, and all that would follow, the students shared their work and were required to justify their decisions in the "why" that they had crafted for the client and their unique user group.

**Figure 4**

*Student Team Activity to Develop Space Planning Options per User Group*

### *Workshop Three*

Workshop three took a break from the activity-based model, and instead students engaged with a multidisciplinary panel of consultants comprised of key design partners in the architectural and interior design field. The panel consisted of an audio-visual theatrical designer, a structural engineer, an electrical engineer, and a mechanical and plumbing engineer. The goal of this panel was to open students' eyes to the reality of design and that all these consultants, framed correctly on a project, are teammates, collaborators, and deeply important to the design process. They were able to discuss how their individual specialties work with and affect interior design, what project timelines look like, the cost, and how important it is to see them all as true partners in the design process and to engage with them early and often. The goal was to expose students to more real-world functions of the design process that they may not have encountered in a traditional studio project and better position them for the day they walk into a firm and collaborate with these specialties.

### *Workshop Four*

The fourth workshop was a hands-on vendor showcase and interior design-focused session. Vendors who represented flooring, lighting, furniture, tile, and fabrics were present with samples appropriate to use in a secondary school environment. Students reviewed their space plans and vision statements before visiting vendor tables to inquire about the best options for their spaces, focusing on more than just aesthetics. As shown in Figure 5, they were encouraged to ask about the durability of products in their specific environment and for their intended user group, as well as about the sustainable features of each item. This not only gave them tactile experiences but also a realistic interaction with vendors who they would be interfacing with in the future. Building on their iterations, they were tasked with sketching furniture layouts into their space plans, crafting elevations or perspectives of key areas, and then taking the finishes they had gathered along with their plans to craft a board showcasing their choices. Groups

presented their designs to simulate client interior finishes and furniture presentation, where they were able to receive instantaneous feedback from industry mentors and the vendors who donated their time and materials to this workshop.

**Figure 5**
*Hands-On Vendor Showcase of Furniture, Fixtures and Finishes*

***Workshop Five***

The final workshop focused on branded environments, with an expert from the industry presenting on how a designer can layer in signage, wayfinding, donor recognition, graphics, and placemaking into a space and how interior designers can collaborate with such a specialty. This final step in the design showcased how this layer of branded environments elevates the interior designer's story to the next level, making an entire experience in a space cohesive. Students were tasked with incorporating everything they had created in the prior workshops, including research, space planning, user needs and materiality, to develop a holistic package for how branded environments could layer into their design for the school. This final workshop took the students back to the big picture of the school as a whole and showed how far they had come in this series. Students had now walked the full arc of a real design process, from research and visioning to detailing and branded storytelling, all guided by working professionals who design such spaces on a daily basis.

## Results

Through the analysis process, the data collected from the survey, interviews and observational field notes were triangulated to confirm emerging themes and reduce the risk that findings reflected a single data method (Maxwell, 2008).

During the workshops, observations by the researchers found that the student participants were instantly engaged in the activity, demonstrated creativity and a willingness to listen to the industry professionals and their peers for new perspectives on the design proposal. Results from

the post-workshop survey revealed that 50% of all participants attended all five workshops, which demonstrates a strong retention rate for a voluntary, co-curricular series spanning two semesters. Common goals students shared for participating in the workshops were to enhance their cross-disciplinary collaboration skills, learn more about educational design, and gain insights into the practices of real-world firms.

To answer the second research question, examining ways industry professionals leading design workshops enhance students' transferable, real-world readiness, particularly empathy and client-centered thinking, the survey revealed that students with little to no knowledge of the educational design industry showed increased interest in this field after attending the workshops. Specifically, 86% of the participants expressed attraction to this specialized career path as a result, and further findings noted that 86% of students said the workshops were helpful in advancing their educational goals. From a student interview, when asked what the student learned, they responded, “it was eye-opening to see how a real design firm handles a real-world project, how they navigate competing ideas and constraints and that getting that insider look ahead of an internship was incredibly valuable.” Confirming that industry mentorship increases knowledge and shifts career trajectories.

The third research question was whether relocating design learning activities outside the traditional studio environment influences student motivation, creativity, and engagement with the design process. The survey found that 71% of students developed a new preference in how they like to learn about interior design. From the open-ended questions in the survey, one student noted, “Because I generally don't enjoy group work, the workshop changed my perspective and made me realize the importance of having multiple perspectives and ideas.” Another shared, “I felt like I was able to understand other people's perspectives as well. It's great to learn from others.” And a third reflected, “I got to see new perspectives on how we can collaborate and create a space.” Input from one of the student interviews revealed that working on design solutions within the vertical collaboration among peers pushed the student to branch outside their normal circle of friends and see a real range of design perspectives and personalities. In summary, the findings indicate that the workshops not only facilitated knowledge content but also shifted students’ understanding and value of the learning process.

In reflection of the first research question, how the combination of PBL and experiential learning impacts engagement and comprehension in interior design education, one comment from a student interview summed up a repeated observation made by the researchers over the course of the workshops. The student expressed how they overcame their fear of speaking up, due to lack of confidence, but the workshops opened their eyes to the idea that no question and no idea is wrong, and instilled confidence in their presentation and communication skills. Such an outcome can be difficult to manufacture in a traditional lecture-and-critique studio model. It requires sustained, repeated, low-stakes exposure to real professionals and real peer collaboration over time.

## Discussion

Tying together industry professionals, cross-disciplinary perspectives, and vertical student integration into one design project is difficult for faculty to replicate in a classroom setting. The results of the project-based, experiential learning workshops demonstrated that professionals bring real-time project knowledge and process to the students, easily replicating complex, client-centered professional practice. Students displayed a high level of engagement in the activities plus a willingness to listen to the professionals. This raises a real institutional

question of how academia could formally recognize, structure, and sustain these industry co-educator roles within curriculum design, rather than treating them as occasional guest visitors.

In addition, student reflections consistently pointed to working alongside education majors, communication designers, and art educators as a transformative part of their experience. Mixing students of different education levels, plus different majors, broadened perspectives, led to peer-to-peer mentorship, leadership development for older students and confidence building for younger students. This invites a broader conversation about whether cross-disciplinary and vertical student collaboration should be formally embedded as a required learning outcome across all years of interior design education.

Given the workshops were conducted as a voluntary extracurricular activity, some students may not have been able to attend due to personal circumstances, such as work. Finding opportunities, such as scaling down the workshops to bring such an experience into the classroom, would create equity of access for all interior design students.

Since these workshops were specifically designed to prepare students for complex, client-centered professional practice, the natural next step in this research would be to assess whether participants in future studio courses or as they enter careers demonstrate a deeper development of empathy for a client, reach across disciplines for design collaboration, or develop a stronger professional performance.

## Conclusion

This study revealed that integrating project-based learning with experiential, industry-led workshops offers an effective model for enhancing design education. It expands learning beyond the studio walls, strengthens professional and peer connections, plus prepares interior design students for the realities of complex, client-centered practice. The qualitative data, triangulated across surveys, interviews and researcher observation, converge on similar themes, with students describing gains in confidence, communication skills and deeper empathy for end users and with 50% of students attending all five workshops, suggests a level of sustained engagement. Worth noting is the 71% participant shift in preferred learning style and 86% career-trajectory shift. The presented research highlights the potential for transformative learning experiences that benefit students by fostering a deeper connection between theoretical knowledge and real-world application and preparing students for career demands.

Given the study's limitations, including findings from one interior design program in the U.S., generalizability across institution types and disciplines is limited. Furthermore, with participation as voluntary, the sample could be skewed toward students already inclined toward collaborative, hands-on learning. Despite these limitations, the research provides a foundation for educators seeking ways to integrate experiential learning with industry professionals meaningfully into design curricula.

## Author's Note

Katie Branning is currently affiliated with HKS, Inc., Dallas, Texas, USA.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author(s) used Grammarly, an AI-assisted writing software, to proofread the manuscript for grammatical accuracy, proper verb tenses, and spelling errors. The author further declares that, apart from Grammarly, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Asojo, A., & Vo, H. (2021). Pedagogy+ Reflection: A problem-based Learning case in interior design. *International Journal of Designs for Learning*, *12*(2), 1–14. https://doi.org/10.14434/ijdl.v12i2.25372

Ayob, A., Majid, R. A., Hussain, A., & Mustaffa, M. M. (2012). Creativity enhancement through experiential learning. *Advances in Natural and Applied Science*, *6*(2), 94–99.

Buchal, R. O. (2011). The educational value of student design competitions. *Proceedings of the Canadian Engineering Education Association* (CEEA). https://doi.org/10.24908/pceea.v0i0.4036

The Buck Institute for Education. (2019). *PBLWorks*. https://www.pblworks.org/

Grossman, P., Pupik Dean, C. G., Kavanagh, S. S., & Herrmann, Z. (2019). Preparing teachers for project-based teaching. *Phi Delta Kappan*, *100*(7), 43–48. https://doi.org/10.1177/0031721719841338

Guo, P., Saab, N., Post, L. S., & Admiraal, W. (2020). A review of project-based learning in higher education: Student outcomes and measures. *International Journal of Educational Research*, *102*, 101586. https://doi.org/10.1016/j.ijer.2020.101586

Institute for Experiential Learning. (2026, July 25). *What is experiential learning?* https://experientiallearninginstitute.org/what-is-experiential-learning/

Knowles, M. S., Holton III, E. F., & Swanson, R. A., & Robinson, (2025). *The adult learner: The definitive classic in adult education and human resource development.* (10th ed.). Routledge.

Kolb, D. A. (1984). *Experiential learning*. Prentice-Hall.

Kolb, D. A. (2015). *Experiential learning: Experience as the source of learning and development.* (2nd ed.). Pearson Education.

Maxwell, J. A. (2008). Designing a qualitative study. In L. Bickman & D. J. Rog (Eds.), *The SAGE handbook of applied social research methods* (2nd ed., pp. 7). SAGE Publications.

Rashdan, W. (2017). Academics and Professionals: Collaborative Strategies for an Interior Design Curriculum. *The International Journal of Design Education, 11*(3), 29-46. https://doi.org/10.18848/2325-128X/CGP/v11i03/29-46

Roberts, J., & Design, C. E. (2007). The future of academic-industry collaboration. *Proceedings of the International Association of Societies of Design Research Conference: Emerging Trends in Design Research*, 1–17. The Hong Kong Polytechnic University School of Design.

Tamim, S. R., & Grant, M. M. (2013). Definitions and uses: Case study of teachers implementing project-based learning. *Interdisciplinary Journal of Problem-Based Learning, 7*(2). https://doi.org/10.7771/1541-5015.1323

Vail, J., & Ellis, N. (2023). Advancing the value of design competitions through experiential learning. In G. Hurcomb, E. Herring, J. Jackson, & V. Peu Duvallon (Eds.), *Teaching Beyond the Curriculum: A Focus on Pedagogy 2023* (AMPS Proceedings Series, Vol. 36.2, pp. 92–100). AMPS. https://amps-research.com/proceedings/

Wobbe, K. K., & Stoddard, E. A. (Eds.). (2023). *Project-based learning in the first year: Beyond all expectations*. Taylor & Francis.

Xi, J., & Wang, X. (2022). Development of landscape architecture design students' pro-environmental awareness by project-based learning. *Sustainability*, *14*(4), 2164. https://doi.org/10.3390/su14042164

**Contact emails:** jeannine.vail@unt.edu
kbranning@hksinc.com

# Towards Immersive Training for Inclusive Education: Co-designing a Virtual Reality Application for Educators

Malak Kanaan, ENSAM, France
Sylvain Fleury, ENSAM, France
Vivien Walter, i-magineXR, France
Karine Galmiche, i-magineXR, France
Simon Richir, ENSAM, France
Nour El Mawas, Université de Lorraine, France

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This study examines how educators' needs can inform immersive training scenarios for supporting learners with Autism Spectrum Disorder (ASD). Grounded in Design-Based Research, it used a sequential mixed-methods design: an online questionnaire-based survey, interviews and co-design workshops. Survey results showed a gap between direct experience and specific training, alongside conditional interest in virtual reality (VR)-based professional development. Interviews identified recurring challenges involving sensory overload, explicit communication, oral assessment, and unexpected transitions. Workshop participants evaluated eight interview-derived scenarios and prioritized three: instruction lost in classroom noise, refusal to give an oral presentation, and an unexpected classroom transition. The study contributes a traceable process linking field needs to refined, prioritized training scenarios and derives design implications for a future immersive educator training.

*Keywords:* inclusive education, Autism Spectrum Disorder, virtual reality, educator professional development, Design-Based Research

iafor
The International Academic Forum
www.iafor.org

## Introduction

Inclusive education requires educators who are both equipped with practical strategies and confident in their ability to support learners with Autism Spectrum Disorder (ASD), yet international guidance continues to identify a persistent gap between the policy commitment to inclusion and the practical preparedness of the educators expected to deliver it (UNESCO, 2017). Immersive technologies, and virtual reality (VR) in particular, are increasingly explored as a way to let trainees rehearse difficult classroom situations from a first-person perspective, with recent systematic evidence suggesting a positive, if heterogeneous, effect of VR on teacher-education outcomes (Billingsley et al., 2019). Realizing this potential for ASD-inclusive training specifically requires deliberate design attention: autistic people are a heterogeneous population, and simulations that are not carefully designed risk flattening that diversity into a single, stereotyped portrayal (Maun et al., 2024; Newbutt & Bradley, 2025), underscoring the need for a solid, field-grounded design process to unlock VR's potential for ASD-inclusive educator training, rather than counting against its use.

This paper reports the needs-analysis phase of a Design-Based Research (DBR) project that aims to design and evaluate a VR application for training educators to support ASD learners. Rather than starting from a technological solution, the project starts from documented classroom experience: an online questionnaire-based survey and semi-structured interviews with educators were used to surface recurring difficulties and existing pedagogical strategies, and two co-design workshops with professionals then used this material to evaluate, refine, and prioritize pre-drafted candidate training scenarios. The paper's contribution is methodological and empirical: it shows how field-level needs were progressively translated into three concrete, prioritized immersive scenarios, and derives design implications for the next, still-ongoing development phase.

## Related Work

### Educator Preparedness for ASD-Inclusive Teaching

International policy frames inclusive education as a systemic transformation rather than an add-on (UNESCO, 2017), yet educators frequently report themselves under-prepared, describing persistent gaps between formal training and "real-world" practice (Devi et al., 2024). This gap is especially salient in higher education: autistic students remain an under-recognized university population whose sensory, communicational, and organizational needs persist or intensify under new academic demands (Anderson et al., 2017). Autistic students' own French accounts of university life echo this pattern, describing sensory overload, inconsistent accommodation, and the effort of seeking help (McPeake et al., 2023), while inclusive-pedagogy practice remains unevenly adopted among university staff themselves (Kennel et al., 2021). In-service training reviews similarly find active, practice-based components (coaching, modelling, feedback) consistently more effective than didactic delivery alone (Wang et al., 2024) – a direct rationale for an experiential, rehearsal-based medium such as VR.

### Sensory Processing and the Case for Immersive Rehearsal

Between 60% and 96% of autistic people report sensory sensitivities, and difficulty filtering competing auditory input is a mechanistically distinct barrier to attention and social processing, not simply distractibility (Mallory & Keehn, 2021). VR and mixed-reality simulation are increasingly used to let teachers and educators rehearse communication and classroom-

management skills under repeatable, low-stakes conditions (Larson et al., 2025), with reported gains in preparedness for challenging behaviour specifically (Billingsley et al., 2019; Chen, 2022; Cumming et al., 2025). Evidence specific to autism shows VR-based perspective-taking can build longer-lasting empathy than traditional exercises (Herrera et al., 2018), and a randomized trial found VR effective for autism-awareness training specifically (Koniou et al., 2025).

## Research Questions

Taken together, the literature highlights both a persistent need for practice based professional development in ASD inclusive education and the potential of immersive simulation to support experiential learning. However, less is known about how educators' field-reported needs and classroom experiences can be systematically translated into immersive training content. Accordingly, this study addresses the following research questions:

RQ1. What training needs and recurring classroom difficulties do educators themselves identify in relation to ASD-inclusive teaching?

RQ2. How can these needs be translated into, and prioritized as, immersive training content?

## Method

### Design-Based Research and Mixed-Methods Design

The study follows a Design-Based Research approach (Sanchez & Monod-Ansaldi, 2015), combining a sequential mixed-methods needs analysis (questionnaire-based survey followed by interviews) with co-design workshops. Each phase built on the previous one: survey responses informed the interview guide and identified volunteers for the interviews, and the interview themes were used to draft the eight candidate scenarios subsequently evaluated in the co-design workshops.

### Questionnaire

An online questionnaire-based survey was made available in French and English versions; respondents completed whichever language version they preferred. It was distributed through professional mailing lists in higher-education based in France, academic social networks (Linkedin), and direct contact. The survey collected demographic information, direct ASD-related experience, prior training, current inclusive practices, and attitudes toward VR and co-design participation. It received 92 responses in total for analysis. The full questionnaire, including an English version, is provided as supplementary material.

### Semi-structured Interviews

Thirteen individual video-call interviews were conducted with 14 participants (one session was a joint interview with two interviewees), recruited primarily from questionnaire-based survey respondents who indicated interest in further participation. All 14 were selected for their direct expertise in supporting ASD learners: each was an experienced practitioner (teacher, teacher-trainer, specialized educator, or pedagogical-support professional) who already had, or had previously had, at least one learner with ASD in their own classroom or caseload. Interviews were semi-structured: each opened with the same core question about the participant's professional background and ASD-related experience, after which further questions probed

classroom experiences, pedagogical strategies, institutional constraints, and attitudes toward immersive training, with question order and follow-up adapted to each participant's responses (full interview guide, including an English version, provided as supplementary material). Interviews were between 45 and 60 minutes. Each session was recorded and transcribed for analysis purposes only; recordings were not retained beyond the analysis or shared publicly.

### Co-design Workshops

Two 2-hour online co-design workshops (8 and 10 December 2025) were conducted as live video-conference sessions (Microsoft Teams) using a shared Miro board to present eight pre-drafted candidate scenarios, each specifying a narrative summary, a point-of-view (learner or teacher), and an intended sensory/interactive mechanic. Eleven distinct professionals participated across the two sessions (6 in the first, 5 in the second; no overlap between the two groups). Ten had also taken part in the interview phase;one had not: a speech-language-therapy student recruited via the facilitator's professional network for the first session. The two sessions used an identical protocol and scenario set and were conducted independently: second session participants were not informed of the first session's discussion or votes, and the scenario materials were not modified between sessions. Participants rated perceived realism, discussed causes and possible teacher responses, and could cast up to three dot-votes to prioritize which scenarios to develop first in VR; 29 votes were placed across the two sessions. Both sessions were recorded and transcribed for analysis purposes only, on the same basis as the interviews; recordings were not retained beyond the analysis or shared publicly.

### Data Analysis

Questionnaire-based survey data were analysed using descriptive statistics (frequencies, percentages, and selected cross-tabulations on the largest valid subgroups), computed directly from the cleaned response dataset; open-ended responses were grouped through thematic content analysis. Interview transcripts were imported into NVivo 14 (Lumivero, 2023) and analysed inductively: familiarization with each transcript, open coding of recurring difficulties and strategies, and consolidation of codes into higher-order themes through iterative comparison across participants. NVivo was used to organize the transcripts, manage the coding structure, and retrieve coded excerpts; coding decisions were made by the researcher, not automated. Workshop analysis combined verbatim session transcripts, written Miro contributions, realism assessments, and vote counts to identify recurring arguments, divergences, proposed adaptations, and scenario-prioritization criteria. The open-text item on current pedagogical practice was additionally visualized as a frequency-scaled phrase cloud (Figure 3) after removing common French stopwords and merging obvious spelling or accent variants. Qualitative coding (interviews and workshops) was conducted by one researcher, without a second coder or inter-rater reliability check.

## Results

Figure 1 previews how the three phases connect to the three prioritized scenarios reported in detail below.

**Figure 1**
*The Three-Phase Research Process, From Field Needs to Prioritized Immersive Scenarios*

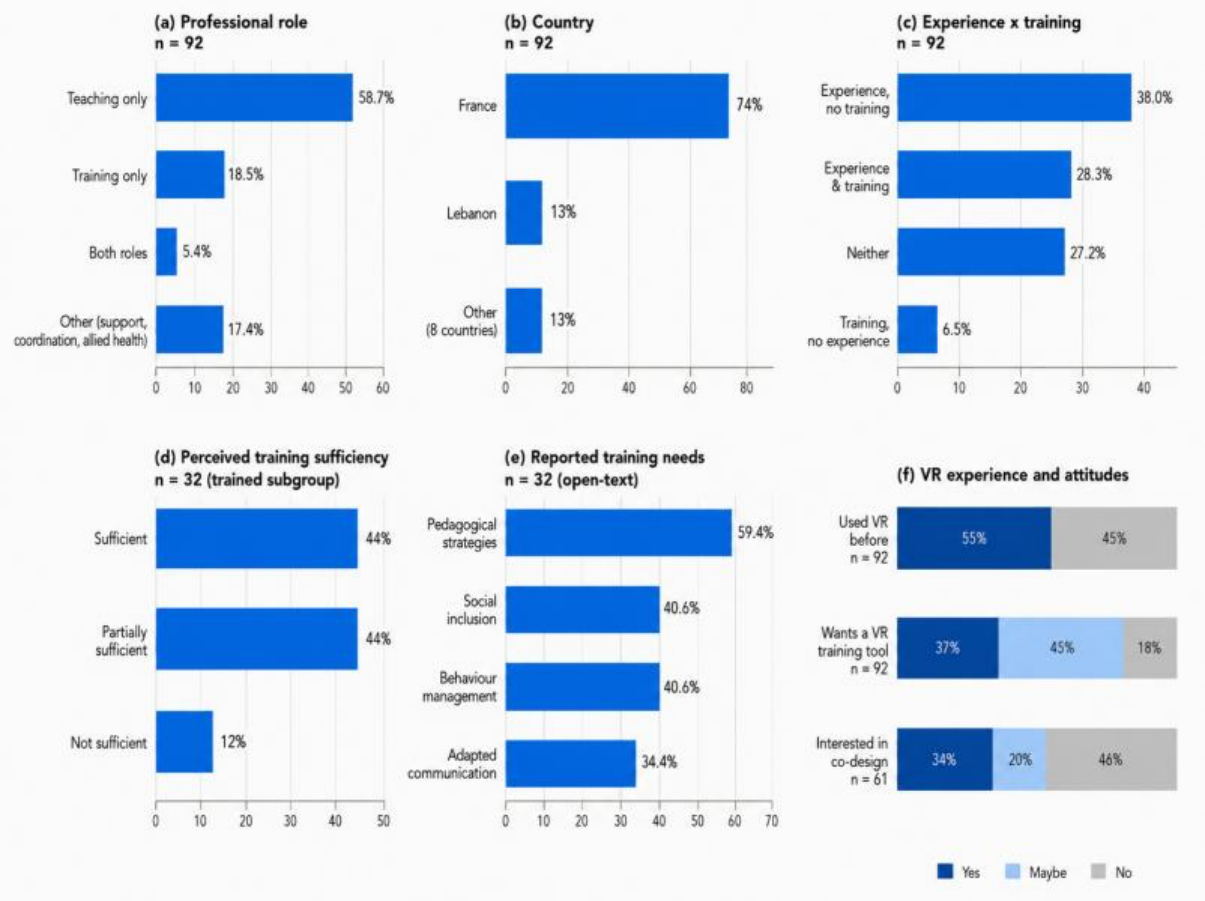
## Questionnaire-Based Survey: Sample and Training Needs (N = 92)

Respondents were 51.1% female, 45.7% male, and 3.3% preferred not to say, with a mean age of 41.7 years (SD = 11.8) and a mean of 14.7 years of professional experience (SD = 10.4). Contrary to a general "school teacher" framing, most respondents reported teaching at university/higher-education level and identified primarily as teachers or teacher-trainers, with a minority in support, coordination, or allied-health roles (Figure 2a). The sample was drawn predominantly from France and Lebanon, with eight further countries (Switzerland, Belgium, Niger, Benin, Germany, Estonia, Spain, Slovakia) each represented by one to four respondents (Figure 2b). Because France and Lebanon differ in composition, we checked whether the overall patterns reported below are geographically confounded: higher-education status varies by country (88.2% of French respondents vs. 58.3% of Lebanese respondents teach at university level), whereas direct ASD experience does not vary as much (France 66.2%, Lebanon 75.0%, other countries 58.3%); these are descriptive comparisons on small national subgroups, not inferential tests, and are reported to make the sample's composition transparent rather than to claim a country effect.

Sixty-six percent had already supported at least one ASD learner, mostly in higher-education or mainstream classroom settings, while only 34.8% had received specific ASD/inclusive-education training (Figure 2c); this asymmetry – 38.0% of respondents had direct experience without specific training, versus only 6.5% the reverse – indicates that direct classroom experience with ASD learners is not, on its own, translating into formal preparation for most respondents. Training itself did not fully close the gap: of the 32 respondents who had received specific training (the same 32 respondents throughout panels 2d and 2e below), only 43.8% rated it fully sufficient, and self-rated ASD knowledge was descriptively higher among this trained subgroup (M = 4.13/5) than the untrained majority (M = 2.78/5, computed from a separate self-rating item asked of all 92 respondents), a descriptive association rather than a tested effect. Among these 32 respondents, the areas they most wanted to deepen were pedagogical strategies and social inclusion (Figure 2e; respondents could select more than one area, so percentages sum to more than 100%). Fifty-five percent had already used a VR device, predominantly for recreational rather than professional or ASD-specific purposes; VR use was itself somewhat more common among French respondents (60.3%) than Lebanese respondents (33.3%). Prior VR use was not associated with stronger interest in a VR training tool (37.3% "yes" among prior VR users vs. 36.6% among non-users), but it was associated with more outright refusal (23.5% "no" among prior users vs. 12.2% among non-users) and less hedging ("maybe": 39.2% vs. 51.2%) – prior exposure to VR appears to sharpen respondents' opinions in both directions rather than simply increasing demand. Among the 61 respondents with direct ASD experience, 34.4% expressed interest in participating in a co-design study (Figure 2f).

**Figure 2**

*Questionnaire Results Overview (N = 92): (a) Professional Role, (B) Country, (C) Direct ASD Experience Crossed With Prior Training, (D) Perceived Sufficiency of That Training Among the Trained Subgroup, (E) Reported Training Needs, and (F) VR Experie*




The most frequently reported strategies for supporting ASD learners were adaptation of teaching materials (adaptation des supports; 17/60), individualized support (accompagnement individuel; 9/60), visual communication (communication visuelle; 8/60), and work in pairs (travail en binôme; 6/60), reflecting a broader emphasis on adapting instructional materials and formats, clarifying communication, and providing individually tailored support (Figure 3). Several respondents also reported exempting ASD students from oral presentations or group activities rather than adapting the activity itself – a real-world precedent directly relevant to one of the co-design workshops' prioritized scenarios.

**Figure 3**

*Frequency-Scaled Phrase Cloud of Pedagogical Strategies Reported by Questionnaire-Based Survey Respondents, in Their Original French Wording (Open-Text Responses, N = 60; Phrase Size Represents the Number of Respondents Who Mentioned Each Expression)*

## Interviews Recurring Themes

Thematic analysis of the 14 interview participants identified eight higher-order themes, presented below together with the scenario each most directly informed.

Sensory overload was the most pervasive concern, spanning two related themes: difficulty filtering a teacher's voice from competing classroom sounds (raised by most participants), and, more broadly, visual, tactile, and multisensory regulation – light, touch, smell, stimming, and the stigma-risk of visible quiet spaces. Several participants, some of whom also drew on lived experience of autism, described their own post-lecture shutdowns in terms that illustrate the felt intensity of this difficulty, not only its frequency: "c'est comme si j'étais dans un concert sans bouchon d'oreille toute la journée" ("it's like being at a concert without earplugs all day long"; anonymized interview excerpt, translated from French) – a first-person account that directly supports modelling this theme as an experiential rehearsal scenario rather than a lecture topic.

A near universal recommendation concerned the use of explicit, sequenced, and dually oral and written instructions. Participants also emphasized the importance of clarifying implicit social expectations and avoiding ambiguous language. Discussions of social inclusion and masking revealed an unresolved tension between respecting a learner's preference to work alone and encouraging participation in group activities. Oral-presentation anxiety emerged as a related concern, particularly when expectations, speaking order, or assessment criteria were unclear. More broadly, participants repeatedly stressed the need to interpret behaviours commonly perceived as challenging as possible signs of anxiety, sensory overload, or unmet needs rather than deliberate defiance. Sudden or insufficiently signposted changes in routine were frequently described as factors that could intensify distress.

Participants also identified institutional and systemic barriers, including difficulties related to disclosing a diagnosis, accommodations being implemented more consistently during examinations than in everyday teaching, and limited participation in autism-awareness training; one participant reported, for instance, that only around 30 of an estimated 800 eligible staff had attended such training at their institution – a self-reported figure that was not independently verified and is reported here as an illustration of perceived low uptake rather than an audited statistic. They also highlighted persistent misconceptions that an awareness-raising tool should address, particularly the tendency to portray autistic people as either exceptionally capable or severely impaired, without recognizing the wide diversity of profiles between these extremes, as well as misconceptions concerning eye contact.

**Co-design Workshops: Scenario Evaluation and Prioritization**

The eight candidate scenarios were derived directly from the interview themes and translated recurring classroom experiences into concrete situations suitable for immersive training. During both workshops, participants assessed their realism, discussed possible causes and pedagogical responses, and cast up to three votes for the scenarios they considered most relevant for future development. Because most workshop participants had also taken part in the interviews from which the scenarios were derived, realism assessments are best read as informed appraisal by an engaged practitioner community rather than fully independent validation; scenarios synthesized recurring patterns across the full interview sample rather than any single participant's account, which limits, without eliminating, the risk of participants simply re-endorsing their own prior statements. Across the 29 votes recorded (14 in the first workshop, 15 in the second), the three highest combined totals were obtained by Instruction Lost in Classroom Noise (4 + 4 = 8 votes), Refusal to Give an Oral Presentation (1 + 4 = 5 votes), and An Unexpected Classroom Transition (2 + 2 = 4 votes). The noise and transition scenarios received stable support in both sessions, whereas the oral-presentation scenario showed greater between-workshop variation (from the least-voted scenario in the first session to tied for most-voted in the second). The three scenarios were retained on this combined prioritization and the pedagogical richness of the discussions they generated; the five non-prioritized scenarios are documented in supplementary materials rather than in the main text, for space.

**Table 1**
*Prioritized Co-design Scenarios: Brief Description and Targeted Educator Competencies*

| Scenario | Brief Description | Educator Competencies for Autism-Inclusive Practice |
|---|---|---|
| **Instruction Lost in Classroom Noise** | An important oral instruction becomes difficult to process because it is delivered amid competing classroom sounds and simultaneous visual or physical activity. | Recognizing sensory overload and auditory-filtering difficulties; reducing simultaneous stimuli; providing concise and multimodal instructions; using visual supports; checking individual understanding without interpreting confusion as inattention. |

| | | |
|---|---|---|
| **Refusal to Give an Oral Presentation** | An autistic learner refuses to give an oral presentation in front of the group, creating tension between common assessment requirements and individual accommodation needs. | Distinguishing anxiety from defiance; applying equitable rather than identical treatment; clarifying expectations and speaking order; proposing equivalent assessment formats; communicating without minimizing the learner's distress. |
| **An Unexpected Classroom Transition** | A learner who has not completed or mentally closed an activity resists an unexpected transition to the next task. | Supporting cognitive flexibility and executive functioning; anticipating changes; using explicit and literal language; providing visual and temporal cues; clarifying how and when an unfinished task will be resumed. |

## Discussion: Integrating the Three Phases

### What Each Phase Contributed

The three phases did not produce three disconnected findings; each narrowed and grounded the next, and together they answer RQ1. The questionnaire-based survey established the breadth of the training gap across the broader sample: two-thirds had direct ASD experience but only a third had received specific training, and over half of that trained minority still felt under-prepared, alongside broad but conditional openness to a VR training tool. The interviews supplied the situated and explanatory detail the survey could not reach: not just that "communication" and "behaviour management" were named needs, but why participants attributed lost instructions to competing sensory channels rather than simply to inattention, and why apparent "defiance" is more productively read as an unmet need. Read together, the survey and interviews show that educators' self-identified training needs (RQ1) are recurring, situated, and only partly addressed by current practice, rather than a diffuse or idiosyncratic list of complaints.

### Why These Scenarios Matter, and What This Study Contributes

The workshops answer RQ2. The three prioritized scenarios are not simply the most frequently mentioned difficulties; they share a structure well suited to immersive rehearsal. Each depicts a brief, high-stakes decision point at which a small, low-cost change in the educator's own behaviour determines whether the situation de-escalates or compounds (for example, pausing before speaking, offering an alternative format, or giving advance warning of a transition): the kind of in-the-moment judgment that lecture-based training struggles to convey but a repeatable simulation can. They also cover complementary mechanisms rather than one repeated theme: sensory filtering, social anxiety, and executive-function/routine change, together representing a deliberately broad first slice of the difficulties surfaced across all three phases rather than a narrow or redundant set. This is an empirical description of what the workshops produced; the pedagogical claims that follow are our interpretation of why it matters, not a further finding.

Beyond the three scenarios themselves, this study's main contribution is methodological: a transparent, traceable progression from field-reported needs to prioritized training content, in which each scenario can be traced back to specific survey findings and interview themes and its inclusion justified by an auditable workshop vote count rather than by design judgment alone. The approach also operationalizes a principle several participants insisted on: that training content should be grounded in, and negotiated with, practitioners who have direct field experience of ASD-inclusive teaching, not designed about them in their absence. This is a standard for how the content was developed, not a requirement that future trainees already share that experience – the resulting scenarios are intended precisely for educators who do not yet have it. This study approximates that co-design standard through educator and researcher collaboration, while explicitly falling short of directly including autistic co-designers, an important direction for subsequent development phases.

## Conclusion

This study set out to identify educators' self-reported training needs and classroom difficulties in ASD-inclusive teaching (RQ1), and to examine how these needs can be translated into, and prioritized as, immersive training content (RQ2). The questionnaire-based survey and interviews answer RQ1: educators across the sample described a persistent gap between direct classroom experience and specific ASD training, and identified recurring, situated difficulties – sensory overload, ambiguous or implicit communication, oral-assessment anxiety, and unsignposted transitions – that current practice addresses only unevenly. The co-design workshops answer RQ2: translating these needs into training content is achievable through a traceable process in which interview themes are drafted into candidate scenarios, then appraised for realism and prioritized by a second group of practitioners, yielding three concrete, evidence-grounded scenarios – an instruction lost in classroom noise, a refusal to give an oral presentation, and an unexpected classroom transition – together with concrete design implications for each. Taken together, these findings are a complete, self-contained contribution: a transparent, auditable path from field-reported needs to prioritized immersive training content, independent of whether a prototype is subsequently built. Developing and technically specifying these scenarios into a working VR prototype, and evaluating its usability, acceptability, and training effect once available, is future work that this study's findings make possible rather than work this study depends on to be complete. The study's scope is also its main boundary: the sample is weighted toward higher-education staff in France and, secondarily, Lebanon, so the difficulties and scenarios reported here should be read as grounded in that context rather than assumed to generalize directly to K-12 settings or to countries outside this sample; extending this needs-analysis-to-scenario process to other education levels and national contexts is a natural direction for future work.

## Ethics and Data Protection

The study received a favourable opinion from the *Comité d'Éthique de la Recherche Arts et Métiers* (reference CER_AM_2025_012, committee session of 4 September 2025). Participation in the questionnaire, interviews, and co-design workshops was voluntary, and participants gave informed consent prior to data collection, with the right to withdraw at any time. Questionnaire respondents could optionally provide an email address to receive information about the continuation of the study or to volunteer for subsequent interviews and co-design activities. These contact details were stored separately from the analytical dataset and were not used in the analysis or reporting of results. Direct identifiers were removed from the data used for analysis; institutions are not named, interview quotations are presented as

anonymized excerpts, and workshop participants are described only by professional role. All data were stored securely and processed in accordance with the approved protocol.

## References

Anderson, A. H., Stephenson, J., & Carter, M. (2017). A systematic literature review of the experiences and supports of students with autism spectrum disorder in post-secondary education. *Research in Autism Spectrum Disorders*, *39*, 33–53. https://doi.org/10.1016/j.rasd.2017.04.002

Billingsley, G., Smith, S., Smith, S., & Meritt, J. (2019). A systematic literature review of using immersive virtual reality technology in teacher education. *Journal of Interactive Learning Research*, *30*(1), 65–90.

Chen, C. Y. (2022). Immersive virtual reality to train preservice teachers in managing students' challenging behaviours: A pilot study. *British Journal of Educational Technology*, *53*(4), 998–1024. https://doi.org/10.1111/bjet.13181

Cumming, T. M., Lee, H. M., & Hoenig, J. (2025). Using mixed reality simulations to prepare preservice teachers for inclusive classroom management. *Australasian Journal of Special and Inclusive Education*, *49*(2), 126–137. https://doi.org/10.1017/jsi.2025.10005

Devi, A., Palmer, E. E., Ganguly, R., & Barua, P. D. (2024). Teachers' educational experiences and preparedness in teaching students with autism. *Asia-Pacific Education Researcher*, *33*(1), 71–81. https://doi.org/10.1007/s40299-022-00709-7

Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLoS ONE*, *13*(10), e0204494. https://doi.org/10.1371/journal.pone.0204494

Kennel, S., Guillon, S., Caublot, M., & Rohmer, O. (2021). La pédagogie inclusive : représentations et pratiques des enseignants à l'université [Inclusive pedagogy: Representations and practices of university teachers]. La Nouvelle Revue – Éducation et Société Inclusives, (3–4), 23–45. https://doi.org/10.3917/nresi.090.0023

Koniou, I., Douard, E., & Lanovaz, M. J. (2025). Brief report: Virtual reality to raise awareness about autism. *Journal of Autism and Developmental Disorders, 55*(9), 3378–3386. https://doi.org/10.1007/s10803-023-06216-y

Larson, K. E., Myers, L., & Hirsch, S. (2025). Using mixed-reality simulation and coaching in special education courses to improve communication skills with paraeducators. *Journal of Special Education Technology.* Advance online publication. https://doi.org/10.1177/01626434251403025

Lumivero. (2023). *NVivo* (Version 14). Computer software.

Mallory, C., & Keehn, B. (2021). Implications of sensory processing and attentional differences associated with autism in academic settings: An integrative review. *Frontiers in Psychiatry*, *12*, Article 695825. https://doi.org/10.3389/fpsyt.2021.695825

Maun, R., Fabri, M., & Trevorrow, P. (2024). Participatory methods to engage autistic people in the design of digital technology: A systematic literature review. *Journal of Autism and Developmental Disorders*, *54*(8), 2960–2971. https://doi.org/10.1007/s10803-023-06015-5

McPeake, E., Lamore, K., Boujut, É., Khoury, J., Pellenq, C., Plumet, M.-H., & Cappe, É. (2023). "I just need a little more support": A thematic analysis of autistic students' experience of university in France. *Research in Autism Spectrum Disorders*, *105*, Article 102172. https://doi.org/10.1016/j.rasd.2023.102172

Newbutt, N., & Bradley, R. (2025). Exploring ethical research issues related to extended reality technologies used with autistic populations. *Journal of Responsible Technology*, *21*, Article 100102. https://doi.org/10.1016/j.jrt.2024.100102

Sanchez, É., & Monod-Ansaldi, R. (2015). Recherche collaborative orientée par la conception. Un paradigme méthodologique pour prendre en compte la complexité des situations d'enseignement-apprentissage [Design-oriented collaborative research: A methodological paradigm for accounting for the complexity of teaching-learning situations]. *Éducation et didactique, 9*(2), 73–94. https://doi.org/10.4000/educationdidactique.2288

UNESCO. (2017). *A guide for ensuring inclusion and equity in education.* UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000248254

Wang, H.-T., Chen, P.-J., Lin, F.-Y., Chen, M.-C., & Hu, S.-T. (2024). Systematic review of effects and key components of in-service special education teacher training in evidence-based practices for students with autism and developmental disabilities. *Education and Training in Autism and Developmental Disabilities*, *59*(1), 97–112. https://doi.org/10.1177/215416472405900109

**Contact email:** malakkanaan17@gmail.com

# A Tale of Two Assessments: Navigating Generative AI Integration in Authentic Assessment Design

Ramalakshmi Vaidhiyanathan, Cardiff University, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

The study presents a comparative analysis of two approaches to authentic assessment design, highlighting how scaffolding and generative AI integration influence student learning in computing education. In the first approach, students engaged in structured coursework supported by predefined requirements and starter code. A custom AI chatbot was provided to enhance comprehension and guide processes while deliberately restricting code generation to uphold academic integrity. This structured design fostered strong student engagement, reduced instructor intervention, and supported consistent understanding of expectations. Although some evidence suggested limited AI-assisted code generation, the controlled scope mitigated the risk of overreliance on automated solutions. In contrast, the second approach adopted an open-ended use of generative AI, permitting students to employ AI tools as virtual clients to create requirement specifications and to use AI-generated code with declaration. This high-autonomy model unintentionally elevated student anxiety and increased requests for instructor support. Students faced difficulties determining appropriate requirements to demonstrate learning outcomes and struggled to connect AI outputs to underlying design principles. Instead of empowering learners, unrestricted AI use resulted in decision paralysis and conceptual confusion. Together, these findings underscore that autonomy without adequate scaffolding can overwhelm learners engaging with emerging technologies. Effective AI integration requires calibrated structure that supports critical thinking, along with explicit instruction on evaluating AI outputs against learning objectives as a core digital literacy competency.

*Keywords:* generative AI, authentic assessment, scaffolding, computing education, student engagement

iafor
The International Academic Forum
www.iafor.org

## Introduction

The integration of generative artificial intelligence (AI) within higher education has prompted renewed and continued scrutiny of assessment design, particularly in computing disciplines where both procedural and conceptual understanding are required (Biscalchin et al., 2025). Recent discussions emphasize the need to focus on the underlying assessment construct, ensuring that tasks accurately measure intended learning outcomes rather than peripheral skills or tool usage (Hao et al., 2024). This study presents a comparative case analysis of two approaches to authentic assessment implemented in a second-year module on object-orientation, algorithms, and data structures.

The study examines how differing levels of scaffolding and autonomy influence student engagement, understanding, and performance when generative AI is incorporated into assessment tasks (Prompiengchai et al., 2025). Specifically, it contrasts a structured assessment model supported by a constrained AI chatbot for support with an open-ended model permitting unrestricted AI usage for both requirement generation and implementation.

By analyzing these approaches, the study reiterates the need for more robust, policy-aware approaches to AI integration in assessment design, including human-centered frameworks that ensure alignment between technological tools and pedagogical intent (Biscalchin et al., 2025). This comparison provides insight into how assessment design choices shape student experience and highlights the importance of aligning AI use with clearly defined learning objectives and assessment constructs (Hao et al., 2024; Prompiengchai et al., 2025).

## Literature Review

The rapid adoption and accessibility of generative artificial intelligence (AI) tools in education have generated significant debate concerning their pedagogical implications, particularly in relation to assessment integrity, student learning, and assessment design (Biscalchin et al., 2025; Hao et al., 2024). Recent work highlights the need to rethink traditional assessment approaches to ensure alignment with learning outcomes in the presence of generative AI (Hao et al., 2024; Khan et al., 2023).

A central concern is ensuring that assessment tasks remain focused on the intended construct (Hao et al., 2024). In AI-integrated environments, the distinction between student-generated and AI-assisted output can become blurred, raising concerns around validity and authenticity (Hao et al., 2024; Khan et al., 2023). Emerging frameworks emphasize the importance of aligning AI use with pedagogical intent through structured and transparent design principles (Biscalchin et al., 2025; Prompiengchai et al., 2025).

The literature also highlights the importance of scaffolding when students engage with complex tasks and emerging technologies. Without sufficient guidance, students may face increased cognitive load and struggle to manage both the conceptual and operational demands of AI-enhanced tasks (Biscalchin et al., 2025).

Furthermore, authentic assessment in the age of AI must shift focus from product to process. Emphasizing how students arrive at solutions rather than solely evaluating outputs has been identified as a key strategy for maintaining academic integrity and supporting meaningful learning (Khan et al., 2023).

The ability to critically evaluate AI-generated outputs is increasingly recognized as a core digital literacy skill. Students must develop the capacity to assess the quality, relevance, and appropriateness of AI outputs in relation to learning objectives (Hao et al., 2024; Vadaparty et al., 2025).

These considerations underpin the present study, which examines how varying levels of scaffolding and AI autonomy shape student experiences in authentic assessment contexts.

## Methodology

This study adopts a comparative case study design, analyzing two iterations of assessment implemented within the same module context.

In the first iteration, students engaged in structured coursework characterized by predefined requirements and the provision of starter code. A custom AI chatbot was integrated to support comprehension and guide task processes (Prompiengchai et al., 2025). Critically, the chatbot was restricted from generating code, thereby functioning as a pedagogical support tool rather than a solution generator.

In the second iteration, a contrasting design was implemented in which generative AI was explicitly incorporated as part of the assessment process. Students were permitted to use AI tools to generate requirement specifications, effectively simulating interactions with virtual clients.

Students were granted full autonomy to determine which requirements to implement and how to demonstrate learning outcomes relating to design principles and patterns. AI-generated code was permitted, subject to mandatory declaration.

This design enables a direct comparison of structured versus open-ended AI integration, facilitating analysis of their respective impacts on engagement, decision-making, conceptual understanding, and the alignment between student outputs and intended learning outcomes.

## Results

The structured AI-assisted assessment demonstrated positive outcomes in both student engagement and learning processes. The chatbot effectively supported students in interpreting requirements and establishing workflows, resulting in a reduction in instructor queries. Assessment outcomes remained consistent with previous cohorts, with an observable shift towards higher classifications.

Although there was some indication of limited AI-assisted code generation, the constrained design mitigated the risk of overreliance. The structured environment facilitated clarity and enabled students to focus on applying conceptual knowledge.

In contrast, the open-ended assessment approach resulted in increased student anxiety and a higher frequency of instructor queries. Students reported difficulty in determining appropriate requirements and ineffectively evidencing learning outcomes.

The absence of predefined structure introduced significant decision-making challenges. Students frequently sought validation and expressed uncertainty regarding scope. Furthermore,

many encountered difficulties in relating AI-generated outputs to underlying design principles, resulting in conceptual misalignment with intended learning outcomes.

## Discussion

The findings underscore the importance of maintaining alignment between assessment design and intended learning outcomes in AI-integrated environments. The structured approach ensured that AI functioned as a supportive tool without displacing core learning objectives, consistent with recommendations that emphasize principled AI integration in pedagogy (Biscalchin et al., 2025).

By contrast, the open-ended approach shifted significant cognitive and metacognitive demands onto students. Learners were required to assess requirements, evaluate AI outputs, and demonstrate conceptual understanding simultaneously (Hao et al., 2024). This led to increased cognitive load, supporting findings that autonomy without adequate scaffolding can hinder learning outcomes (Biscalchin et al., 2025).

The study also highlights the importance of clearly defining the pedagogical role of AI. Positioning AI as an assistive tool rather than a generative solution appears to promote deeper engagement with learning processes. In contrast, unrestricted AI use may shift focus towards output production, thereby weakening the validity of assessment (Hao et al., 2024).

Additionally, the results reinforce the importance of developing students' critical AI literacy. Students must be able to interrogate and evaluate AI-generated outputs rather than accept them uncritically. This aligns with emerging views that digital and prompt literacy are essential competencies in AI-enabled education (Hao et al., 2024; Prompiengchai et al., 2025).

Overall, the findings suggest that effective AI integration requires careful calibration between structure and autonomy, ensuring that technological affordances support rather than undermine pedagogical goals.

## Conclusion

This study provides a comparative analysis of two approaches to integrating generative AI in authentic assessment design demonstrating that the level of scaffolding significantly influences student experience and learning outcomes.

The structured, scaffolded approach supported student engagement, reduced instructional demands, and facilitated consistent demonstration of learning outcomes. In contrast the open-ended approach characterized by high autonomy and unrestricted AI use, resulted in increased anxiety, decision-making challenges, and difficulties in linking outputs to conceptual understanding.

These findings indicate that the effectiveness of generative AI in education is contingent on the manner of its integration. Structured guidance, clearly defined roles for AI, and explicit support for evaluating AI-generated outputs are essential to ensuring meaningful learning.

The study emphasizes the importance of balancing autonomy with scaffolding in the design of AI-integrated assessments particularly in contexts requiring both technical proficiency and conceptual understanding.

Future assessment design must prioritize structured AI integration, emphasize process-oriented evaluation, and explicitly develop students' ability to critically engage with AI as part of their core disciplinary and digital literacy competencies.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Microsoft Copilot was used in proofreading and refining the language used in the manuscript for accuracy and clarity. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Biscalchin, A., Sarkheyli-Hägele, A., Eriksson, J., & Vogel, B. (2025). Designing Education for the AI Era: Principles for Integrating LLMs in Pedagogy. *2025 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 1–6. https://doi.org/10.23919/softcom66362.2025.11197458

Hao, J., von Davier, A. A., Yaneva, V., Lottridge, S., von Davier, M., & Harris, D. J. (2024). Transforming Assessment: The Impacts and Implications of Large Language Models and Generative AI. *Educational Measurement*, *43*(2). https://doi.org/10.1111/emip.12602

Khan, M. M., Dong, Y., & Manesh, N. A. (2023). Authentic Assessment Design for Meeting the Challenges of Generative Artificial Intelligence. *IEEE Frontiers in Education Conference (FIE)*. https://doi.org/10.1109/fie58773.2023.10343376

Prompiengchai, S., Narreddy, C., & Joordens, S. (2025). *A Practical Guide for Supporting Formative Assessment and Feedback Using Generative AI*. ArXiv.org. https://arxiv.org/abs/2505.23405

Vadaparty, A., Smith IV, D., Srinath, S., Padala, M., Alvarado, C., Benario, J. G., Zingaro, D., & Porter, L. (2025). *Integrating Large Language Models and Evaluating Student Outcomes in an Introductory Computer Science Course*. ArXiv.org. https://arxiv.org/abs/2510.18806

**Contact email:** vaidhiyanathanr@cardiff.ac.uk

# Does More Coursework Mean Better Preparation? Examining Course Availability and Opportunities to Learn Inclusive Practices in Teacher Education

Melike Yozgatli Gulum, Ministry of National Education, Türkiye

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Inclusive education is a central goal of teacher preparation programs, yet questions remain about how program structures shape opportunities to learn inclusive practices. Prior research has emphasized the importance of offering inclusion-related coursework; however, less is known about whether increasing the number of such courses meaningfully supports the integration of inclusive practices. This study examines the relationship between course availability and opportunities to learn inclusive practices in teacher education programs. Using survey data collected from 215 faculty members in special education and elementary education–related programs at four-year higher education institutions in the United States, this quantitative study explores associations between the number of inclusion-related courses offered and faculty-reported integration of inclusive practices. Faculty reported an average of five inclusion-related courses within their programs. While most participants indicated that inclusion is considered when selecting clinical or field placement opportunities, fewer than two-thirds reported that all pre-service teachers received clinical or field placements in inclusive settings. Analyses revealed weak but statistically significant associations between course availability and overall integration of inclusive practices. Course availability was more strongly associated with collaboration and assessment-related practices, whereas relationships with instructional and social–emotional–behavioral practices were weak and not statistically significant. These findings suggest that increasing the number of inclusion-related courses alone has limited explanatory power for how inclusive practices are integrated in teacher preparation. The results highlight the importance of intentional course design and practice-based and field-based opportunities in shaping meaningful opportunities to learn inclusive practices.

*Keywords:* teacher education, inclusive practices, course availability, opportunity to learn

iafor
The International Academic Forum
www.iafor.org

## Introduction

Inclusive education has become an increasingly prominent goal of educational systems, reflecting efforts to ensure that students with disabilities have meaningful access to general education settings alongside their peers (Artiles & Kozleski, 2007; Downing & Peckham-Hardin, 2007; Guldberg, 2010). In inclusive classrooms, however, placement alone does not ensure successful educational experiences. Teachers must be prepared to respond to diverse academic, behavioral, social, and emotional needs through effective and evidence-based practices. Consequently, teachers play a central role in determining the quality of inclusive education, as they are responsible for translating inclusive policies and principles into daily classroom practice (Martínez, 2003; McLeskey et al., 2018).

The growing presence of students with disabilities in general education classrooms has increased expectations for both general and special education teachers. Teachers are expected not only to deliver instruction aligned with the general education curriculum but also to provide individualized supports, use assessment information to guide instruction, address social-emotional and behavioral needs, and collaborate effectively with families and other professionals (Pit-ten Cate et al., 2018). Yet previous research has consistently raised concerns about teachers' preparedness to meet these expectations. Teachers and pre-service teachers have reported limited confidence and preparation for implementing inclusive practices, often attributing these difficulties to insufficient training and limited opportunities to develop the knowledge and skills needed in inclusive settings (Robinson, 2017; Woodcock & Hemmings, 2012).

Teacher education programs therefore have an important responsibility to prepare pre-service teachers for the realities of inclusive classrooms. Research has demonstrated the potential value of coursework, field experiences, collaborative activities, and practice-based preparation in improving teachers' knowledge, attitudes, and perceived ability to support students with disabilities (Alexiadou & Essex, 2016; Goddard & Evans, 2018; Sharma & Nuttal, 2016). Pre-service teacher education has also been identified as an important context for developing positive attitudes toward inclusion and increasing teachers' self-efficacy for implementing effective practices in inclusive settings (Goddard & Evans, 2018; Lambe & Bones, 2007).

One common response to concerns about teacher preparedness has been to increase the availability of coursework related to special education, inclusion, collaboration, and instructional support for students with disabilities. Prior research has examined whether teacher preparation programs offer courses addressing inclusion and related content (Bradshaw & Mundia, 2006; Burke & Sutherland, 2004; Sharma & Nuttal, 2016). For example, studies have documented introductory special education courses, collaboration-focused coursework, and courses addressing the characteristics and instructional needs of students with disabilities. At the same time, the literature has identified considerable variation across programs. Some programs provide opportunities related to inclusion and collaboration, whereas others offer comparatively limited coursework or provide these opportunities primarily to special education majors (Allday et al., 2013; Harvey et al., 2010; McKenzie, 2009).

The availability of inclusion-related coursework, however, represents only one dimension of teacher preparation. Counting the number of courses offered provides limited information about what pre-service teachers actually have opportunities to learn and do within those courses (McKenzie, 2009). Inclusive practices may be introduced through lectures or readings, assessed through examinations, applied through projects and lesson planning, or practiced in

authentic field-based settings. These experiences represent substantially different levels and forms of engagement with inclusive practices. Previous concerns about teacher preparedness suggest that exposure to inclusive education content without sufficient opportunities for application and practice may not adequately prepare future teachers for the complexity of inclusive classrooms. Indeed, shortcomings in teacher preparation have frequently been linked not simply to the absence of inclusive content but also to insufficient practice-based preparation and limited opportunities to develop inclusive teaching skills.

High-leverage practices provide a useful framework for examining the preparation of pre-service teachers for inclusive education. Organized around collaboration, assessment, social-emotional-behavioral practices, and instruction, these practices represent important competencies for supporting students with disabilities in general education classrooms (McLeskey et al., 2017, 2018).

Although increasing the number of inclusion-related courses may provide additional exposure to these practices, it remains unclear whether course availability itself corresponds to greater integration of inclusive practices across teacher education programs. The distinction between quantity of coursework and the depth of opportunities to learn within coursework is therefore important. Programs may offer multiple inclusion-related courses while still providing relatively limited opportunities for pre-service teachers to apply, demonstrate, and receive feedback on inclusive practices. Conversely, programs with fewer courses may integrate these practices more intentionally and extensively within existing coursework. Accordingly, the present study examined the relationship between the number of inclusion-related courses offered and faculty-reported integration of inclusive practices in teacher education programs in the United States. Specifically, the study asked whether the number of inclusion-related courses was associated with the overall integration of inclusive practices and with integration within four domains: collaboration, assessment, social-emotional-behavioral practices, and instruction.

## Literature Review

### Inclusive Practices

Inclusive education requires teachers to use a range of practices that address the academic, social, emotional, and behavioral needs of students with disabilities in general education classrooms. Because teachers are responsible for implementing these practices in daily instruction (Martínez, 2003), teacher preparation programs play an important role in ensuring that pre-service teachers develop the knowledge and skills needed to work effectively in inclusive settings. Research has emphasized that inadequate preparation and limited experience with inclusive practices may contribute to teachers' difficulties in meeting the diverse needs of students with disabilities (Alexiadou & Essex, 2016; Blanton et al., 2011; McLeskey et al., 2018; Pit-ten Cate et al., 2018; Robinson, 2017; Woodcock & Hemmings, 2012).

One framework for identifying practices that are particularly important for teacher preparation is the set of high-leverage practices developed for special education. These practices were designed to represent research-supported practices that can be implemented across instructional settings and student needs and that can be taught and practiced by pre-service and in-service teachers. They are organized into four domains: collaboration, assessment, social-emotional-behavioral practices, and instruction (McLeskey et al., 2017). The framework has also been

proposed as a potential core component of practice-based teacher preparation (McLeskey et al., 2018).

Collaboration is a central component of inclusive education because supporting students with disabilities often requires coordinated work among general and special education teachers, families, paraprofessionals, and other professionals. Effective collaboration includes working with other professionals to improve student outcomes, conducting productive meetings with families and professionals, and collaborating with families to ensure that students receive appropriate services and support (McLeskey et al., 2018). Previous research has nevertheless identified limitations in how teacher preparation programs prepare pre-service teachers for collaborative practices such as co-teaching and interdisciplinary work (Botha & Kourkoutas, 2016; Chitiyo & Brinda, 2018; Waitoller & Artiles, 2013).

Assessment practices are similarly important because effective inclusive teaching depends on understanding students' strengths and areas of need and using assessment information to make instructional decisions. These practices include gathering information from multiple sources, using assessment results to develop and implement educational plans, and monitoring student data to adjust instruction (McLeskey et al., 2018). Research has also suggested that pre-service teachers may have misconceptions about assessment and may require greater opportunities to develop competence in using assessment information to guide their teaching (Lin & Lin, 2015, 2019).

Social-emotional-behavioral practices focus on establishing supportive learning environments and addressing students' social, emotional, and behavioral needs. These practices include creating positive and well-organized classroom environments, providing constructive feedback, teaching social behaviors and skills, and developing behavioral support plans (McLeskey et al., 2018). Their inclusion in teacher preparation is important because simply placing students with disabilities in general education classrooms does not ensure positive social, emotional, or behavioral outcomes. Teachers must be prepared to provide the supports necessary for students to participate meaningfully in inclusive settings (McLeskey et al., 2018).

Instructional practices represent the largest domain of the high-leverage practices framework. They include identifying learning goals, systematically designing instruction, adapting curriculum and materials, using explicit instruction, employing flexible grouping strategies, supporting student engagement, using technology, and promoting maintenance and generalization of learning (McLeskey et al., 2018). These practices require teachers to combine pedagogical and content knowledge with an understanding of individual student needs. Limited training and experience with inclusive instruction may therefore create challenges when teachers begin working with students with disabilities in general education classrooms.

Together, these four domains provide a framework for considering what pre-service teachers should have opportunities to learn during teacher preparation. However, the presence of inclusive practices within teacher education coursework does not necessarily indicate the extent or depth to which pre-service teachers engage with those practices.

## Opportunities to Learn Inclusive Practices

Teacher education programs can provide opportunities to learn inclusive practices in substantially different ways. A practice may be introduced through lecture or required reading, assessed through a quiz or examination, applied through assignments such as lesson planning

or projects, or practiced and evaluated in field-based settings. Consequently, identifying whether inclusive content is present in a program provides only limited information about the nature of the preparation pre-service teachers receive.

In the present study, opportunities to learn inclusive practices were conceptualized through five ways in which faculty could integrate a practice into their coursework: lecture, required readings, assessment, application assignments, and field-based assignments. Application assignments included lesson plans, demonstrations, projects, and reflections, whereas field-based assignments involved opportunities such as practicum, fieldwork, tutoring, and student teaching. This approach distinguishes simple exposure to inclusive content from opportunities to apply and demonstrate inclusive practices.

This distinction is particularly relevant because inclusive practices are inherently applied. For example, collaboration requires interaction with other professionals and families, assessment requires interpretation and use of student information, and instructional and social-emotional-behavioral practices require teachers to make decisions in response to students' individual needs. Although lectures and readings can provide important conceptual knowledge, opportunities to apply practices and receive feedback may be necessary for pre-service teachers to translate that knowledge into classroom practice.

Previous research has highlighted the potential value of inclusive coursework, training in inclusive settings, and collaborative experiences for strengthening pre-service teachers' knowledge, attitudes, and perceived efficacy regarding inclusive education (Alexiadou & Essex, 2016; Goddard & Evans, 2018; Sharma & Nuttal, 2016). At the same time, concerns remain regarding insufficient practice-based preparation and teachers' confidence in implementing inclusive practices (Pit-ten Cate et al., 2018; Robinson, 2017; Woodcock & Hemmings, 2012).

These concerns suggest that evaluating teacher preparation solely by the number of inclusion-related courses offered may provide an incomplete picture. Offering more courses may increase pre-service teachers' exposure to inclusive education content, but it does not necessarily indicate that those courses provide deeper opportunities to apply, practice, and receive feedback on inclusive practices. Conversely, programs with fewer inclusion-related courses may still provide substantial preparation if inclusive practices are intentionally embedded across existing coursework and field experiences.

The distinction between course quantity and opportunities to learn is therefore important when examining teacher preparation for inclusive education. Rather than assuming that a greater number of courses necessarily represents stronger preparation, it is necessary to examine whether course availability is associated with greater integration of inclusive practices. This distinction provides the basis for the present study's examination of the relationship between the number of inclusion-related courses offered and faculty-reported integration of inclusive practices across collaboration, assessment, social-emotional-behavioral, and instructional domains.

# Methodology

## Participants and Procedures

This study used a quantitative, self-report survey design to examine how inclusive practices are integrated into teacher education coursework in the United States. The target population consisted of faculty members in special education and elementary education–related programs who taught or had previously taught courses addressing inclusive education or related content. Participants were recruited from four-year institutions of higher education that offered undergraduate special education or related programs as well as elementary education or related programs. A convenience sampling approach was used to construct a broad sampling frame because institutional websites did not consistently identify which faculty members taught inclusion-related coursework.

Because institutional websites did not consistently identify which faculty members taught coursework related to inclusive education, contact information was collected broadly from special education and elementary education faculty across eligible institutions. Potential participants were identified through university websites and other higher education resources, and faculty members whose contact information was publicly available were invited to participate electronically.

The survey was administered electronically using Qualtrics following approval from the University of Kentucky Institutional Review Board. Participants received an invitation describing the purpose of the study and were asked to confirm that they taught or had taught coursework related to inclusive education. Two reminder emails were distributed after the initial invitation. Following data screening and cleaning, the final analytic sample consisted of 215 faculty members.

Participants represented special education, elementary education, and combined special and elementary education programs. Of the 215 participants, 96 identified as faculty in special education programs, 40 in elementary education programs, and 77 in combined programs; two did not specify their program affiliation. Participants had an average of 13.13 years of experience in teacher preparation, with experience ranging from one to 46 years. The institutions represented all four major geographic regions of the United States.

## Instrument

Data were collected using a researcher-developed survey designed to measure the extent to which inclusive practices were integrated into teacher education coursework. The survey was developed because no existing instrument was identified that directly measured the integration of inclusive or high-leverage practices within teacher preparation coursework.

The instrument contained 37 questions organized into three sections: program-related questions, course-related questions, and participant demographics. The program-related section gathered information about characteristics such as program type, number of inclusion-related courses offered, clinical and field placements, and evaluation of pre-service teachers' use of inclusive practices. The course-related section contained 23 inclusive practices adopted from the high-leverage practices framework and organized into four domains: collaboration, assessment, social-emotional-behavioral practices, and instruction (McLeskey, 2018).

For each of the 23 practices, faculty first indicated whether the practice was addressed in their coursework. When a practice was addressed, participants reported the ways in which it was incorporated. Five forms of integration were examined: lecture, required readings, assessment through quizzes or examinations, application assignments, and field-based assignments. Application assignments included activities such as lesson planning, demonstrations, projects, and reflections, whereas field-based assignments included observations of teaching during practicum, fieldwork, tutoring, or student teaching. Participants could select multiple forms of integration for each practice.

The survey was pilot tested with faculty members who represented the target population. Feedback was used to revise wording, clarify terminology, provide definitions, and modify the structure of questions concerning the ways practices were incorporated into coursework.

### Variables and Data Analysis

The primary independent variable for the present analysis was the number of inclusion-related courses offered within participants' teacher education programs. The dependent variable was faculty-reported integration of inclusive practices. Integration was examined both overall and separately within the four domains of collaboration, assessment, social-emotional-behavioral practices, and instruction. The integration measure reflected the number of ways in which practices were incorporated into coursework, ranging from zero to five.

Data were exported from Qualtrics, cleaned, and analyzed using SPSS. Descriptive statistics were used to summarize participant and program characteristics. To address the research question examined in this paper, Pearson correlation coefficients were calculated to determine the relationships between the number of inclusion-related courses offered and the integration of inclusive practices overall and within each of the four domains.

Prior to conducting the correlation analyses, assumptions of normality and linearity were examined. The original course-count variable demonstrated nonlinearity; therefore, a natural logarithmic transformation (LN) of the number of inclusion-related courses was used. The transformed variable showed improved linearity and normality and was used in the subsequent Pearson correlation analyses.

## Results

### Program Characteristics and Opportunities to Learn

Participants reported an average of approximately five inclusion-related courses within their teacher education programs. Most faculty members (72.6%, $n$ = 156) indicated that inclusion was considered when selecting clinical or field-placement opportunities for pre-service teachers. However, only 55.8% reported that all pre-service teachers in their programs received clinical or field placements in inclusive settings. In addition, 85% of participants indicated that pre-service teachers were evaluated on their use of inclusive practices during clinical experiences.

Inclusive practices were generally well represented in coursework. Across the 23 practices examined, faculty reported addressing a mean of 20.71 practices ($SD$ = 3.70), corresponding to approximately 90% of the practices included in the survey. However, the ways in which these practices were integrated varied. Across domains, lecture, required readings, and application

assignments were used more frequently, whereas assessments and field-based assignments were used less frequently.

This pattern was evident across the four inclusive-practice domains. Collaboration practices were typically addressed through lectures, readings, and application assignments, with assessments and field-based assignments used less often. Assessment practices were highly represented in coursework but showed a similar pattern, with lecture, readings, and application assignments occurring more frequently than assessment and field-based opportunities. Social-emotional-behavioral and instructional practices likewise tended to be incorporated primarily through lectures, readings, and application assignments rather than through assessments or field-based assignments.

### Relationship Between Course Availability and Integration of Inclusive Practices

Pearson correlation analyses were conducted using the natural-log-transformed number of inclusion-related courses. The number of courses offered was positively and statistically significantly associated with the overall integration of inclusive practices, $r = .151$, $p = .027$. The magnitude of the association was weak.

At the domain level, the number of inclusion-related courses was positively and significantly associated with the integration of collaboration practices, $r = .181$, $p = .008$, and assessment practices, $r = .181$, $p = .008$. Both relationships were weak.

In contrast, the relationships between course availability and the integration of social-emotional-behavioral practices, $r = .090$, $p = .186$, and instructional practices, $r = .118$, $p = .084$, were positive but not statistically significant.

**Table 1**
*Correlations Between the Number of Inclusion-Related Courses Offered and Integration of Inclusive Practices*

| *Domain* | *r* | *p* |
|---|---|---|
| Integration of overall inclusive practices | .151* | .027 |
| Integration of collaboration practices | .181** | .008 |
| Integration of assessment practices | .181** | .008 |
| Integration of social-emotional-behavioral practices | .090 | .186 |
| Integration of instruction practices | .118 | .084 |

*Note.* *. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Overall, the correlation analyses indicated that greater course availability was associated with greater integration of inclusive practices and within the collaboration and assessment domains. However, the associations were weak, and the relationships with social-emotional-behavioral and instructional practices were positive but not statistically significant.

## Discussion

The purpose of this study was to examine whether the number of inclusion-related courses offered in teacher education programs was associated with greater integration of inclusive practices. Overall, the findings suggest that course availability is related to the integration of inclusive practices, but the relationship is limited in magnitude and varies across domains. The

number of inclusion-related courses was positively and significantly associated with overall integration of inclusive practices; however, the association was weak ($r$ = .151). Similarly, weak but statistically significant relationships were found for collaboration and assessment practices, whereas relationships with social-emotional-behavioral and instructional practices were not statistically significant.

These findings are important because the availability of inclusion-related coursework has frequently been discussed as an important component of preparing pre-service teachers for inclusive classrooms. Previous research has suggested that limited inclusion-related coursework and insufficient preparation may contribute to teachers feeling unprepared to meet the needs of students with disabilities (Woodcock & Hemmings, 2012). Other studies have emphasized the potential value of specifically designed inclusion-related courses, content-focused special education coursework, training in inclusive settings, and collaborative experiences (Alexiadou & Essex, 2016; Bradshaw & Mundia, 2006; Burke & Sutherland, 2004; Goddard & Evans, 2018; Sharma & Nuttal, 2016).

The present findings provide some support for the importance of course availability, but they also indicate that simply offering more courses is unlikely to explain substantial variation in how inclusive practices are integrated into teacher preparation. The weak overall relationship between course availability and integration is particularly notable given that inclusive practices were already widely represented in participants' coursework. Faculty reported integrating approximately 90% of the 23 inclusive practices examined in the survey. Thus, differences among programs may not be adequately understood by considering only whether inclusion-related content is present or how many related courses are offered. Instead, greater attention may be needed to the ways in which inclusive practices are incorporated within those courses and the opportunities pre-service teachers have to engage with them.

The descriptive findings reinforce this interpretation. Although inclusive practices were frequently included in coursework, faculty most often reported addressing them through lectures, readings, and application assignments. Assessments and field-based assignments were used less frequently across the four domains. These patterns suggest an important distinction between the presence of inclusive content and the depth of opportunities available to learn and demonstrate inclusive practices. Exposure through lecture and reading may contribute to conceptual knowledge, but practices such as collaborating with professionals, interpreting assessment data, supporting student behavior, and adapting instruction also require opportunities for application in authentic or practice-oriented contexts.

The significant relationship between course availability and collaboration practices may reflect the extent to which collaboration-related content can be distributed across multiple inclusion-related courses. Collaboration is a central component of inclusive education and requires pre-service teachers to develop skills for working with general and special educators, families, and other professionals. Previous research has identified rich content and training in collaborative activities as important components of teacher preparation (Goddard & Evans, 2018; Sharma & Nuttal, 2016). At the same time, the association found in the present study remained weak, suggesting that increasing the number of courses alone may not ensure substantially greater opportunities to develop collaborative competence.

A similar pattern was found for assessment practices. Course availability was positively and significantly related to the integration of assessment practices, but the relationship was also weak. Assessment is a particularly important component of inclusive teaching because teachers

must be able to gather, interpret, and use information to identify student needs and make instructional decisions. Although assessment practices were highly integrated into the programs represented in this study, faculty relied more frequently on lecture, reading, and application assignments than on more practice-oriented assessment and field-based experiences. This finding suggests that the number of courses may provide additional opportunities for assessment-related content to appear in the curriculum, but course quantity does not necessarily indicate the extent to which pre-service teachers are required to apply assessment practices in realistic settings.

In contrast, the number of inclusion-related courses was not significantly associated with the integration of social-emotional-behavioral or instructional practices. These findings are especially informative because both domains represent substantial responsibilities for teachers working in inclusive classrooms. Social-emotional-behavioral practices require teachers to create supportive learning environments, provide constructive feedback, teach social skills, and address behavioral needs. Instructional practices include a broad set of competencies related to designing, adapting, and delivering instruction for students with disabilities. The absence of statistically significant relationships in these domains suggests that increasing course availability does not necessarily correspond to greater integration across all areas of inclusive practice.

Taken together, these findings highlight an important distinction between the quantity of inclusion-related coursework and the depth and intentionality with which inclusive practices are addressed within coursework. Strengthening teacher preparation may therefore require programs to examine not only how many inclusion-related courses they offer but also what pre-service teachers are expected to do within those courses. Opportunities to apply skills, demonstrate practices, receive feedback, and participate in field-based experiences may provide a more meaningful indication of the extent to which inclusive practices are embedded in teacher preparation than course counts alone.

The findings therefore do not suggest that inclusion-related coursework is unimportant. Rather, they indicate that more coursework is not necessarily equivalent to substantially greater integration of inclusive practices. Course availability appears to matter to some extent, particularly for collaboration and assessment, but the generally weak relationships suggest that the structure and depth of learning opportunities within coursework deserve greater attention. Teacher education programs may benefit from examining how inclusive practices are intentionally embedded across coursework and how frequently pre-service teachers are provided with opportunities to move beyond exposure toward application and field-based practice.

## Conclusion

The present study examined the relationship between the number of inclusion-related courses offered and the integration of inclusive practices in teacher education programs. The findings showed a statistically significant but weak positive relationship between course availability and the overall integration of inclusive practices. Similar weak positive relationships were found for collaboration and assessment practices, whereas the relationships with social-emotional-behavioral and instructional practices were not statistically significant. Taken together, these findings indicate that offering a greater number of inclusion-related courses may contribute to opportunities to address inclusive practices, but course quantity alone does not appear to

account for substantial differences in how these practices are integrated into teacher preparation.

The findings also draw attention to the nature of the learning opportunities provided within coursework. Although inclusive practices were widely represented across the programs in this study, faculty relied more frequently on lectures, readings, and application assignments than on assessments and field-based assignments. Thus, evaluating teacher preparation solely according to the number of inclusion-related courses may overlook meaningful differences in the ways inclusive practices are incorporated into coursework. Programs may benefit from examining not only whether inclusive content is included, but also whether pre-service teachers have opportunities to apply practices, demonstrate their skills, receive feedback, and use those practices in authentic educational settings.

**Limitations**

Several limitations should be considered when interpreting these findings. The study relied on faculty self-report rather than direct observations of coursework or pre-service teacher performance. In addition, the survey had a response rate of 10.3%, which may limit the generalizability of the findings, although the sample size was sufficient for the analyses conducted in the larger study. The number of inclusion-related courses also represents only one characteristic of teacher preparation and does not directly capture differences in course quality, sequencing, content, or depth. Other program characteristics may contribute to the integration of inclusive practices. Finally, because the analyses were correlational, the findings should not be interpreted as evidence that increasing the number of courses causes greater integration of inclusive practices.

Despite these limitations, the findings have implications for the design of teacher preparation programs. Rather than focusing primarily on adding inclusion-related courses, programs should consider how intentionally inclusive practices are embedded within existing coursework and clinical experiences. Greater emphasis on assessment, application, and particularly field-based opportunities may provide pre-service teachers with opportunities to move beyond exposure toward practicing and demonstrating inclusive practices. This is consistent with the study's broader implication that teacher preparation programs should provide more opportunities for inclusive practices to be addressed through assessment and field-based experiences and should strengthen access to inclusive clinical placements. Ultimately, these findings suggest that the critical issue may be less the number of inclusion-related courses offered than the opportunities those courses provide for pre-service teachers to learn, apply, and practice inclusive teaching.

## Acknowledgments

This paper is based on data collected as part of the author's doctoral dissertation, *Inclusive Practices in Teacher Education Programs* (Yozgatli, 2022), completed at the University of Kentucky. The author gratefully acknowledges the Stuart D. B. Picken Scholarship provided by IAFOR, which supported her participation in the conference.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that ChatGPT (OpenAI), an AI-assisted tool, was used for proofreading and language editing, as English is not the author's first language.

## References

Alexiadou, N., & Essex, J. (2016). Teacher education for inclusive practice: Responding to policy. *European Journal of Teacher Education*, *39*(1), 5–19. https://doi.org/10.1080/02619768.2015.1031338

Allday, R. A., Neilsen-Gatti, S., & Hudson, T. M. (2013). Preparation for inclusion in teacher education pre-service curricula. *Teacher Education and Special Education, 36*(4), 298–311. https://doi.org/10.1177/0888406413497485

Artiles, A. J., & Kozleski, E. B. (2007). Beyond convictions: Interrogating culture, history, and power in inclusive education. *Language Arts, 84*(4), 357–364.

Blanton, L. P., Pugach, M. C., & Florian, L. (2011). *Preparing general education teachers to improve outcomes for students with disabilities*. American Association of Colleges for Teacher Education & National Center for Learning and Disabilities. https://www.ncld.org/wp-content/uploads/2014/11/aacte_ncld_recommendation.pdf

Botha, J., & Kourkoutas, E. (2016). A community of practice as an inclusive model to support children with social, emotional and behavioural difficulties in school contexts. *International Journal of Inclusive Education*, *20*(7), 784–799. https://doi.org/10.1080/13603116.2015.1111448

Bradshaw, L., & Mundia, L. (2006). Attitudes and concerns about inclusive education: Bruneian in-service and preservice teachers. *International Journal of Special Education*, *21*, 35–41.

Burke, K., & Sutherland, C. (2004). Attitudes toward inclusion: Knowledge vs. experience. *Education*, *125*(2), 163–172.

Chitiyo, J., & Brinda, W. (2018). Teacher preparedness in the use of co-teaching in inclusive classrooms. *Support for Learning*, *33*(1), 38–51. https://doi.org/10.1111/1467-9604.12190

Downing, J. E., & Peckham-Hardin, K. D. (2007). Inclusive education: What makes it a good education for students with moderate to severe disabilities? *Research and Practice for Persons with Severe Disabilities*, *32*(1), 16–30. https://doi.org/10.2511/rpsd.32.1.16

Goddard, C., & Evans, D. (2018). Primary pre-service teachers' attitudes towards inclusion across the training years. *Australian Journal of Teacher Education, 43*(6), 122–142. http://dx.doi.org/10.14221/ajte.2018v43n6.8

Guldberg, K. (2010). Educating children on the autism spectrum: Preconditions for inclusion and notions of 'best autism practice' in the early years. *British Journal of Special Education*, *37*(4), 168–174. https://doi.org/10.1111/j.1467-8578.2010.00482.x

Harvey, M. W., Yssel, N., Bauserman, A. D., & Merbler, J. B. (2010). Preservice teacher preparation for inclusion. *Remedial and Special Education*, *31*(1), 24–33. https://doi.org/10.1177/0741932508324397

Lambe, J., & Bones, R. (2007). The effect of school-based practice on student teachers' attitudes towards inclusive education in Northern Ireland. *Journal of Education for Teaching*, *33*(1), 99–113. https://doi.org/10.1080/02607470601098369

Lin, P. Y., & Lin, Y. C. (2015). Identifying Canadian teacher candidates' needs for training in the use of inclusive classroom assessment. *International Journal of Inclusive Education*, *19*(8), 771–786. https://doi.org/10.1080/13603116.2014.970669

Lin, P. Y., & Lin, Y. C. (2019). Understanding how teachers practise inclusive classroom assessment. *Studies in Educational Evaluation, 63,* 113–121. https://doi.org/10.1016/j.stueduc.2019.08.002

Martínez, R. S. (2003). Impact of a graduate class on attitudes toward inclusion, perceived teaching efficacy and knowledge about adapting instruction for children with disabilities in inclusive settings. *Teacher Development*. https://doi.org/10.1080/13664530300200215

McKenzie, R. G. (2009). A national survey of pre-service preparation for collaboration. *Teacher Education and Special Education, 32*(4), 379–393. https://doi.org/10.1177/0888406409346241

McLeskey, J., Barringer, M.-D., Billingsley, B., Brownell, M., Jackson, D., Kennedy, M., Lewis, T., Maheady, L., Rodriguez, J., Scheeler, M. C., Winn, J., & Ziegler, D. (2017). *High-leverage practices in special education*. Council for Exceptional Children & CEEDAR Center.

McLeskey, J., Billingsley, B., & Ziegler, D. (2018). Using high-leverage practices in teacher preparation to reduce the research-to-practice gap in inclusive settings. *Australasian Journal of Special and Inclusive Education*, *42*(01), 3–16. https://doi.org/10.1017/jsi.2018.3

McLeskey, J., Maheady, L., Billingsley, B., Brownell, M., & Lewis, T. (2018). *High leverage practices for inclusive classrooms*. Routledge. https://doi.org/10.4324/9781315176093

Pit-ten Cate, I., Markova, M., Krischler, M., & Krolak-Schwerdt, S. (2018). Promoting inclusive education: The role of teachers' competence and attitudes. *Insights into Learning Disabilities, 15,* 49–63. https://files.eric.ed.gov/fulltext/EJ1182863.pdf

Robinson, D. (2017). Effective inclusive teacher education for special educational needs and disabilities: Some more thoughts on the way forward. *Teaching and Teacher Education*, *61*, 164–178. https://doi.org/10.1016/j.tate.2016.09.007

Sharma, U., & Nuttal, A. (2016). The impact of training on pre-service teacher attitudes, concerns, and efficacy towards inclusion. *Asia-Pacific Journal of Teacher Education*, *44*(2), 142–155. https://doi.org/10.1080/1359866X.2015.1081672

Waitoller, F. R., & Artiles, A. J. (2013). A decade of professional development research for inclusive education: A critical review and notes for a research program. *Review of Educational Research*, *83*(3), 319–356. https://doi.org/10.3102/0034654313483905

Woodcock, S., & Hemmings, B. (2012). Does study of an inclusive education subject influence pre-service teachers' concerns and self-efficacy about inclusion? *Australian Journal of Teacher Education*, *37*(6), 1–11. https://doi.org/10.14221/ajte.2012v37n6.5

Yozgatli, M. (2022). *Inclusive practices in teacher education programs* [Doctoral dissertation, University of Kentucky]. https://doi.org/10.13023/etd.2022.295

**Contact email:** melike.yozgatli.g@gmail.com

# An Exploration of the Use of Musical Activities to Promote Foundational Literacy Competencies for Children Aged 3 to 5 Years

Maria Kay, University of Aberdeen, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

Reports from the USA and the UK suggest that they are facing a literacy crisis. Whilst much research focuses upon the teaching of the alphabetic principle as a starting point for literacy success, more important are the foundational competencies, which children bring to their formal literacy learning. It is these that determine a child's ease of entry to the educational system. Research to date suggests a strong relationship between music and language and that musical activities are advantageous to literacy learning, yet music cannot teach reading. Through a pragmatic, sociocultural lens, this research used a micro-ethnographic, action research study of nine children, aged three to five years, daily, over a ten-week period, to explore how a literacy-focused musical programme can be designed and delivered to promote foundational literacy competencies and experiences. Data from video recordings were analysed through hybrid coding. The findings revealed that when engaging musical activities which promote auditory, motor, cognitive, language and timing skills, with elements of dialogic teaching are integrated in a Literacy-through-Music programme, focusing on explicit literacy outcomes such as language, and phonological and prosodic awareness, children develop skills which will support literacy learning. Observation of the activities also offered the presenter a panoptic view to possibly identify children at risk of reading failure. Highlighting the features of an integrated Literacy-through-Music programme provides a unique contribution to current literature and shows how such a programme can obviate the need for transfer from musical to literacy skills.

*Keywords*: skills gap, early literacy, literacy foundations, preschool literacy, music and literacy

iafor
The International Academic Forum
www.iafor.org

## Introduction

Literacy, the ability to understand and use spoken language and to read and write at a level which enables a person to engage fully in the functions of daily life, is a necessity; but the Annual Literacy Survey in the UK (NLT, 2026) reported that just over a third of children said that they enjoyed reading in their spare time, whilst only 1 in 4 enjoyed writing. Cabrera et al. (2024) reported that 25% of Year 7 pupils (age 11–12) had a reading age below 11, 1 in 5 primary pupils are failing to meet expected reading standards and that 5.1 million adults in England are functionally illiterate. In America the NAEP (National Assessment of Educational Progress) (U.S. Department of Education et al., 2025) reported that only 43% of 4$^{th}$ graders were at a proficient or above, level of reading. Those most at risk included students of colour, from low-income backgrounds, multi-lingual learners and those with disabilities. Roughly a third of twelfth-grade students in 2024 (U.S. Department of Education et al., 2025) were below basic (able to identify relevant details in a text for literal comprehension); a third were at basic level and a third were proficient (connect details across texts and draw complex inferences about author's purpose, tone, word choice and related ideas) or above.

Despite changing methods of teaching over time, some children to continue to struggle with literacy learning. Literacy models such as those of Scarborough's "Reading Rope" (2001) and the Active View of Reading (Duke and Cartwright (2021) evolved from the Simple View of Reading, proposed by Gough and Tunmer (1986), the "Writing Rope", (Sedita, 2019) and the Not-so-Simple View of Writing (Berninger & Winn, 2006,) which evolved from the Simple View of Writing (Berninger et al., 2002), serve to guide literacy instruction.

The reading and writing "ropes" identify the knowledge and skills which contribute to literacy success and provide metaphorical illustrations of how knowledge and skills must be woven together.

**Figure 1**
*Strands of Early Literacy Development (Scarborough, 2001)*

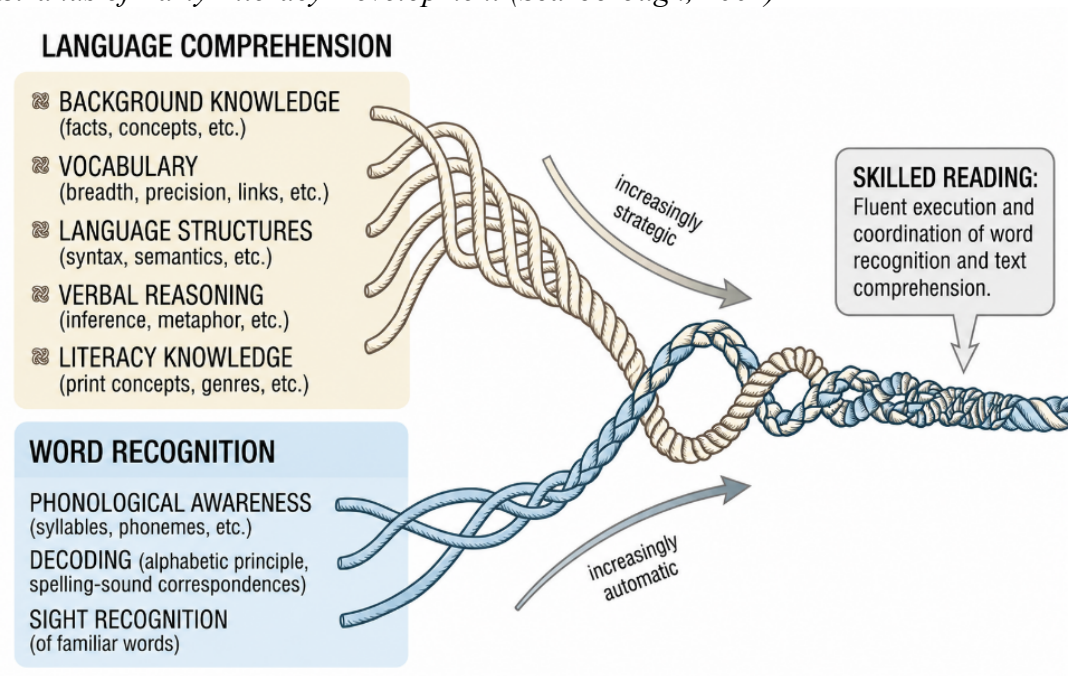
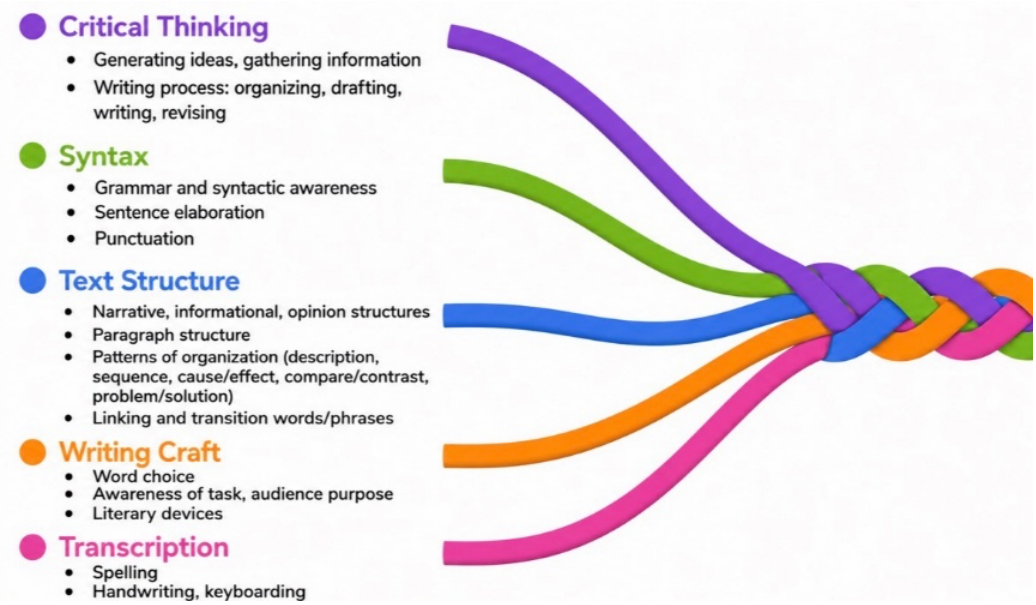


**Figure 2**
*The Strands That Are Woven Into Skilled Writing (Sedita, 2019)*



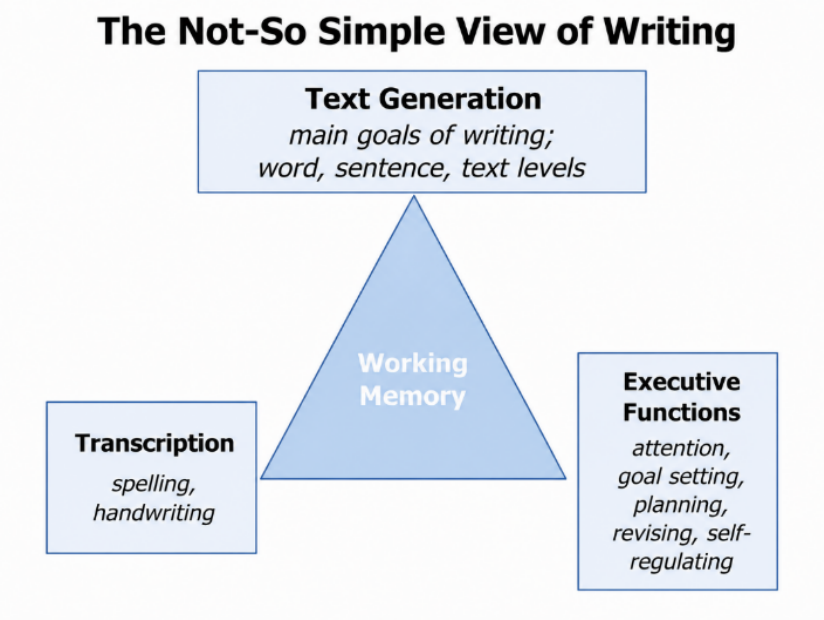

The Active View of Reading recognises the importance of additional skills of active self-regulation and also the impact of the demands of the text, the purpose of reading and the reader's sociocultural background, as these influence how well the reader will comprehend the text.

**Figure 3**
*The Active View of Reading (Duke & Cartwright, 2021)*

Whilst much research focuses upon the skill of reading, reading and writing are complementary skills and the learning of each, supports the learning of the other. As can be seen in Figure 2, the productive task of writing, necessitates the conflation of a great many skills. The Not-so-

Simple View of Writing (Figure 4) acknowledges that composition and transcription are underpinned by executive functions and working memory.

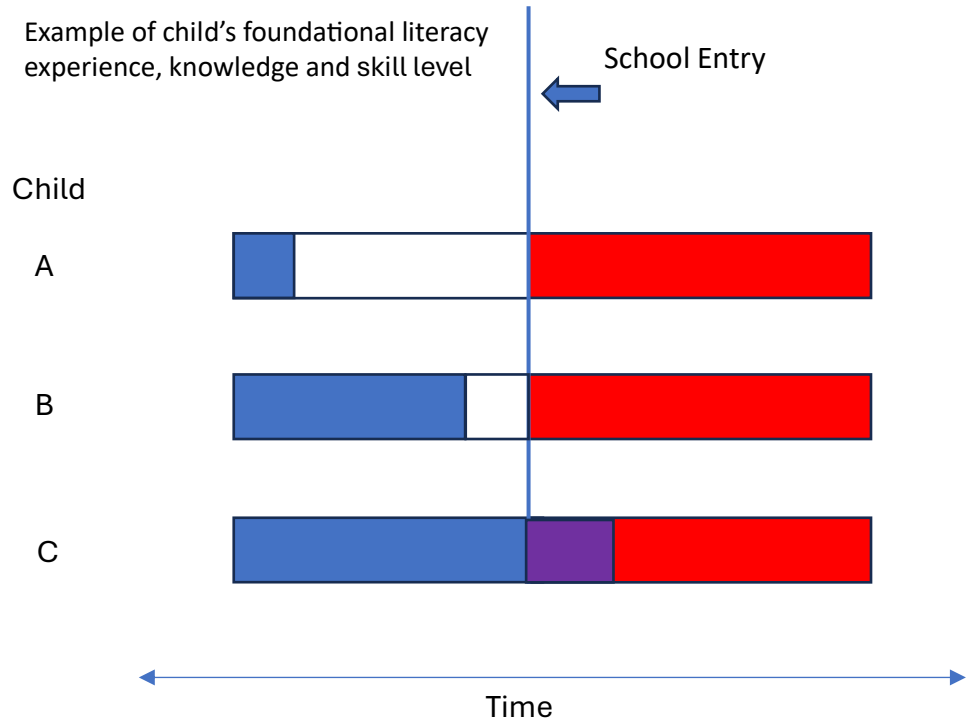
**Figure 4**
*The Not-So-Simple View of Writing (Berninger & Winn, 2006)*




For early years' practitioners there is little to guide foundational instruction, and little qualitative research in this area. Mirhosseini (2017) stated that qualitative research continues to be of lower visibility in the area of language and of literacy education than quantitative research. This research aims to address both identified gaps.

Prior to school entry children acquire knowledge, skills and experience which they will bring to later literacy learning. Rohde (2015) suggested that the informal preschool literacy learning should move seamlessly into formal school literacy learning but this is not the case as children arrive at school with varying competencies. Whilst some children are "ready" to learn to read and write, others are not. For some children the home learning environment can be a rich source of foundational literacy learning (Zhakim, 2024), but for others it is not. Additionally, some children attend preschool settings and others do not. Corriveau et al. (2010) found that children who entered school with little pre-reading skills, phonological awareness and poor receptive vocabulary, were disadvantaged.

**Figure 5**
*Gaps in Early Literacy Learning: Possible Variance in Foundational Literacy Competencies Upon School Entry (Kay, 2026a)*




*Note.* Blue = Foundational literacy experience, knowledge and skills; Red = The teaching of formal literacy skills – phonics; Purple = The teaching of formal literacy skills to a child who already has some of these skills; White = Foundational literacy experience, knowledge and skills gap

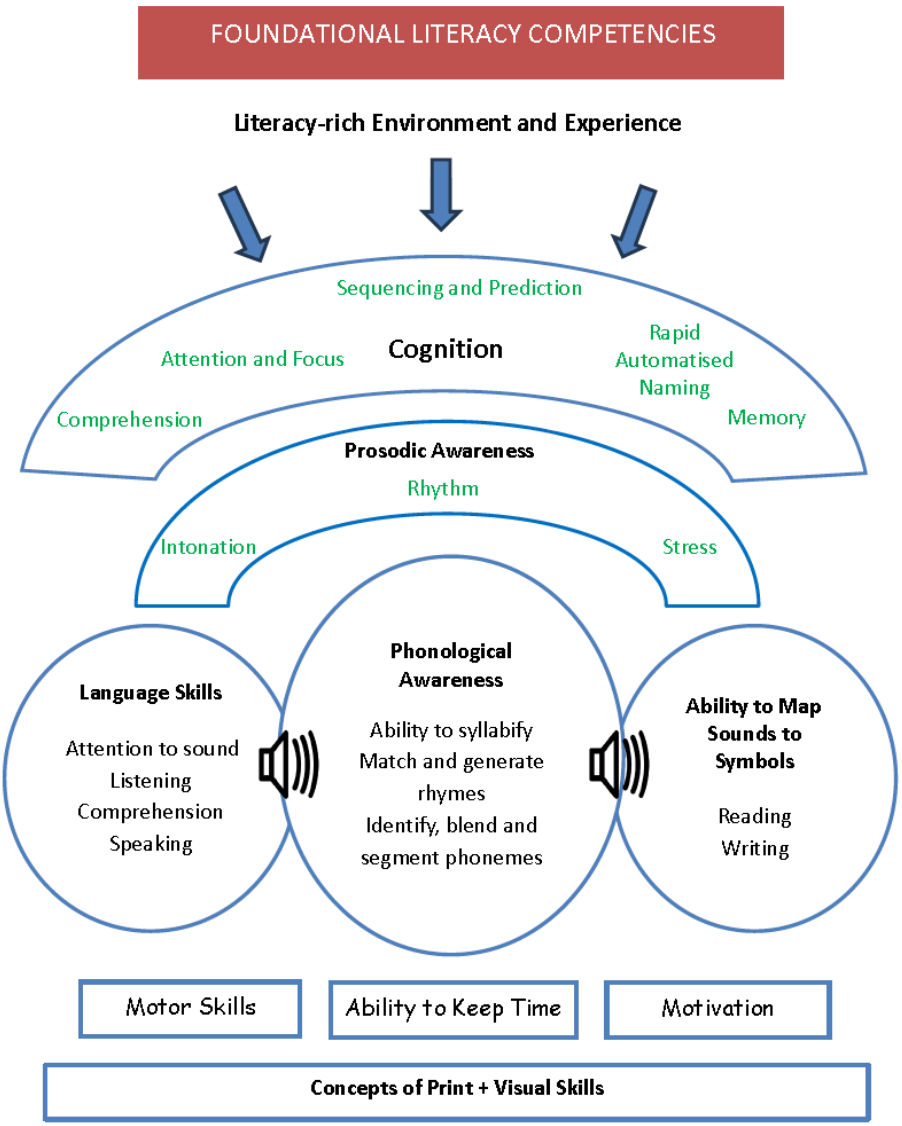
As is evident from Figure 5, the gap, identified in white, illustrates the extent to which some children may not have sufficient literacy experience, knowledge or skills to enable them to be successful upon school entry. Other children may be already reading and writing to some extent.

## Literature Review

Rohde (2015) suggested that when children enter pre-school provision, there is often little understanding of what children require in order to facilitate literacy success and a knowledge of letters is often viewed as the starting point.

A review of the literature revealed the experiences, skills and knowledge upon which literacy success is dependent and which are in place prior to formal literacy learning for those who appear to easily master what is required and which are lacking in those who struggle (Kay, 2026a). These are identified in Figure 6.

**Figure 6**
*Foundational Literacy Competencies (Kay, 2026a)*




When children have been exposed to a literacy-rich environment and have acquired competency is a range of the identified skills, there is a greater chance of literacy success for them than for children without such experiences.

The importance of cognitive skills is often overlooked, when these are important contributors to literacy proficiency. Stothard and Hulme (1992) proposed that comprehension skills are required, Dice and Schwanenflugel (2012) emphasised the importance of being able to focus and pay attention to what is being taught, Stein (2022), purported that being able to sequence and predict what is to follow, supports reading acquisition, and being able to remember both visually and auditorily, what is being taught is essential. McWeeny et al., (2022) found that random automatised naming is predictive of reading ability and thus is also a contributor to literacy proficiency.

Facility with spoken language is imperative to literacy, as written language is predicated upon it. It is important to be able to comprehend and use language and that vocabulary is built over time, as at two years of age, language development is predictive of future literacy performance

(Roulstone et al., 2011). A lack of language skills is a preventative barrier to literacy learning. Visual skills are also acknowledged as important to the recognition and formation of text and it is important to be able to differentiate letter shapes for correct identification.

The ability to identify prosody in language; the stress and intonation placed upon syllables, supports language comprehension and children with dyslexia show poor processing of these stresses (Goswami, 2022). It is imperative that children are able to tune in to and discriminate language sounds in for them to ultimately map them to alphabetic characters. This awareness of sounds within language is termed "phonological awareness" (PA), of which there are three levels—syllable, rhyme and phoneme. Martin et al. (2024) state that PA is the undisputed major predictor of literacy success and that phonemic awareness (awareness of the smallest sound elements within language–phonemes) being the strongest predictor. This ability to recognise phonemes, facilitates their blending and segmenting in order to code and encode written script.

There is a strong positive relationship between motor proficiency and pre-reading skills (Milne et al., 2018). Movement stimulates the brain, thus impacting learning, and gesture supports language learning (Novack & Goldin-Meadow, 2015). Macrine and Fugate (2021, p. 3) refer to the fact that "Our current educational delivery systems and approaches can be traced back to 'disembodied views' of human thinking". This alludes to the general gap between what is known about how children learn – that learning should be active, and the often observation of classroom practice, where they are often observed to be immobile. Macrine and Fugate (2022) advocate for a holistic approach to learning and teaching and assert that embodied learning, where movement supports learning, can promote vocabulary acquisition, language development and comprehension, and handwriting proficiency.

The ability to keep time by entrainment to a beat, encourages fluency of auditory-motor timing, a skill required for reading and writing fluency (Tierney & Kraus, 2014) and as with prosody, there is a strong link between difficulties with rhythm perception and dyslexia. Polyanskaya et al. (2019, p. 1) describe rhythm in speech as "the structure that determines how the signal is organised and develops over time". The reason for the link between reading ability and rhythm memory is elucidated by Tierney et al. (2021) who determined that they rely on shared cognitive processes (such as those used to process abstract concepts, thoughts and memories) and the ability to perceive the temporal structure of sound.

When words are mispronounced, due to ineffective speech processing, it becomes very difficult to match the sounds within the words to corresponding graphemes. As speech rhythm patterns are fundamental to linguistic processing it behoves parents and educators to include infant-directed speech, lullabies and rhymes to enhance language development (Goswami, 2022).

The benefits of participation in musical activities are far-reaching, yet much undervalued (Collins et al., 2020) and include brain stimulation (Schlaug et al., 1995), speech processing (Patel, 2014), assistance with memory recall (Parbery-Clark et al., 2009), encouragement of movement (Toyka & Freund, 2007), enhanced detection of "speech-in-noise" (Slater et al., 2015) and support of the ability to entrain to a beat by engaging auditory-motor systems (Corriveau et al., 2010; Slater et al., 2013). Tierney and Kraus (2013a, p. 228) stated that "one of the reasons musical training can be such a powerful educational tool is that music is inherently rewarding, emotion-inducing and attention grabbing".

In order to learn to read and write, children must be motivated to do so. Rhythm processing and reward have been linked by Fiveash et al. (2023), thus showing how the pleasure gained

through participation in musical activities can be harnessed for learning. Music stimulates emotional areas of the brain and can facilitate memory recall. Where music is integral to learning, the wealth of benefits derived from music alone serve to support learning in other areas and specifically in literacy where music and language share multitudinous commonalities, not least, neural pathways in the brain (Meyer et al., 2014).

Whilst there is a wealth of literature attesting to the correlation between various musical experiences and literacy outcomes, there is no causal evidence to prove that early participation in musical activities results in later literacy success. Reasons for the transfer of skills between music and literacy are suggested by The OPERA hypothesis (Patel, 2011) and the PAT hypothesis, (Tierney & Kraus, 2014), but this cannot be assumed and is not automatic. However, a causal relationship does exist between informal music activities and auditory discrimination (Putkinen et al., 2012) and it is firmly established that auditory discrimination is foundational to phonemic awareness.

The integration of music and literacy activities to promote literacy outcomes would therefore obviate the need for transfer and this has been suggested by Frasher (2014) and Verney (2011), for example. Rautenberg (2015) called for further research in this area. Bolduc and Lefebvre (2012) and Bostelman (2008) suggested that integrating music with literacy learning to promote literacy outcomes have been the most efficacious. The seminal work of Standley and Hughes in 1997, offers one such study and this was replicated by Register in 2001. The results from their study demonstrated that "the more focused the music activities were on a specific skill, the more effective they were in teaching print conventions and writing skills as intended" (Standley & Hughes, 1997, p. 5).

## Method

The micro-ethnographic study was conducted through the analysis of child participant vocalisations and behaviours which were video recorded in their own preschool setting for 30-40-minute periods over six weeks. A setting was chosen in an area of social deprivation, based upon findings from Sosu and Ellis (2014) that children in such areas are at risk of low language abilities which impact upon literacy. There were nine participants, five male and four female, aged between three and five years of age. The participants were chosen randomly by the Centre Head.

A literacy intervention was designed using the "embedded-explicit model" of Justice and Kaderavek (2004), a preventative framework for conceptualising the delivery of emergent literacy interventions for young children who were at risk of later reading difficulties. Literacy outcomes were embedded in the musical activities and were explicitly promoted by the presenter who was a literacy specialist. The study sought to examine how literacy competencies could be promoted through a purpose-designed, integrated literacy-through-music (LtM) intervention.

Activities and props were selected which would stimulate imagination and be engaging for the participants. Music was recorded with introductory and ending sections to help the children to focus upon the musical sounds which indicated when singing should begin and the activity would finish. These sections also allowed participants to anticipate and prepare to sing or perform actions and to anticipate when an activity was coming to an end. The recorded music was syllabic with notes corresponding to the syllables in the lyrics. Whilst the activities, songs, rhymes and movement to music, were musical, their focus was literacy oriented. Initially, focus

was placed upon variations in musical sounds, concordant with a musical program, such as the recognition of rhythm, duration, tempo, dynamics, pitch and timbre. As the program progressed the presenter began to focus on literacy outcomes as identified in Figure 6. Unlike a musical program, the LtM program used books as props to promote bibliographic knowledge, response to visual stimuli, language comprehension and the purpose of text. Dialogic teaching was applied to elicit feedback on literacy outcomes and to promote focus and attention. Towards the end of the program, the presenter used an alphabet book with songs to promote phonemic awareness and the relationship between language sounds (phonemes) and alphabetic symbols (graphemes). All activities were age-appropriate, and letters were presented in a context that would help to embed learning.

### Data Collection and Analysis

The research was conducted according to SERA ethical guidelines (SERA, 2005). Permission slips were obtained from parents and guardians for participation of the children. The children were asked if they wanted to participate and were advised daily that they could leave at any point by asking a member of their staff. At least two centre staff accompanied every session. To ensure the comfort of participants, the sessions were video recorded by volunteer school pupils from the adjoining school. The recordings from twenty-two sessions were suitable for review.

Footage was selected from six full sessions between the beginning to the end of the intervention. These showed the most activities and number of participants. This footage was transcribed. A priori deductive coding was initially used to analyse the transcription and inductive coding was later applied to enable themes to be identified.

## Findings

The data collected provided a wealth of information with regards to foundational literacy skill acquisition through the integrated LtM activities. This enabled a LtM model to be developed. Analysis of the data highlighted the similarities and differences between a musical program and a LtM program in terms of design and delivery. It also enabled specific activities to be identified which could be used to best promote literacy outcomes and highlighted the way in which the presenter was able to do this. The findings revealed how such an intervention could be used as a diagnostic tool to identify strengths and weaknesses in foundational literacy skills.

It also showed how literacy competencies were acquired and positively compounded throughout the program. Six underlying mechanisms were able to be identified to explain how and why the LtM intervention was successful.

At the beginning of the sessions, participants followed actions very easily and were eager to join in. The activities were simple and repetitive, encouraging participation. Music, lyrics and actions were presented simultaneously. Actions were followed more easily than melody or lyrics. Over time, the number and frequency of actions increased, and the length of songs and rhymes increased. Participation then varied between the children as strengths and weaknesses become increasingly visible, the rate of progression and improvement in skill areas differed. However, all participants improved in skill building over time, as was expected. Engagement with the activities was high. Memory retention and recall was demonstrated as participants began to learn melodies and lyrics in addition to actions. Speed of prediction varied between participants. Dialogic teaching was used to assess the comprehension of storylines with

questions being posed and discussion encouraged. For example, the book, "My Cat Ben" (Kay, 2026b) was used and the participants sang the song about a greedy cat, as the music was played, and the pages turned of the book with each verse. As the children became familiar with the book, they were able to remember the order of the verses and to describe the cat and the events in the story. They related these to their own experiences and enjoyed interjecting with their own stories. From the transcription and observation of the video footage, it was possible to observe motivation and demonstration of cognitive and language skills. Awareness and use of prosody was demonstrated. The children were highly engaged and keen to proffer answers to questions posed. Verbal participation increased with confidence over time. Some participants used gesture to support their speech and two used gestures instead of speech on occasions. Using gestures in place of speech indicated immaturity in language development. One participant had a speech impediment, yet a high level of engagement with the activities and frequent contributions and one demonstrated weakness in memory recall.

Participants were shown how to tap out syllables. In one song, the word "alligator" was elongated with the musical notes changing for each syllable. This, along with scaffolding by the presenter, supported all participants to be able to syllabify this word and many others. Children were also able to choose and syllabify words on a theme, themselves, which they all managed successfully by the end of the program.

Participants became adept at matching and generating rhyming words. The strategy of cloze was used by the presenter to encourage the interposing of rhyming words. Attention was drawn to similar sounding patterns at the end of words. One participant was seen to be saying "*cheesey wheezy"* to herself when the presenter spoke about a rhyme with a mouse eating cheese. Although this was not part of the program, her behaviour demonstrated her awareness of and ability to generate rhymes.

Six of the nine participants were notably moving their bodies in time to the music subconsciously. Whilst the perception of beat is said to be innate (Winkler et al., 2009) the ability to synchronise to a beat is not. Carr et al. (2014) have correlated inability to synchronise to a beat with poor literacy skills. Improvement in synchronising to a beat was noted over time in all participants, as was improvement in motor skills.

As knowledge indicators, participants were asked to identify written numbers, name colours and to identify left and right sides of their bodies. Those with weak knowledge were found to have overall weaknesses in all areas. Participants who showed strengths, showed them in all areas.

An unexpected finding was that of volitional rehearsal. Children were reported by staff and parents to be performing some of the activities from the program at other times, other than the scheduled sessions. This also took place within the session, unrequested; for example, one participant was reading aloud printed numbers to himself, which were displayed on the wall. As part of an activity the children had been asked to read aloud the numbers but in this instance the boy was performing of his own accord. Such rehearsal would undoubtedly serve to further embed learning.

When a child was observed to be struggling, scaffolding was implemented by the presenter. The taking of immediate and avertive action could mitigate the risk of future failures in competency assimilation.

## Underlying Mechanisms

Inductive codes were generated to identify the underlying mechanisms for the acquisition of foundational literacy skills. Some activities were observed to be more popular than others. This was due to observed higher levels of participation, they evoked more positive body language (laughing and smiling) and had the greater number of requests for performance or repetition. One particular activity, where children stood at imaginary bus stops and boarded an imaginary bus when their name was sung, was particularly popular. These activities were also the ones for which most skills were demonstrated. Ultimately six possible underlying mechanisms were identified:

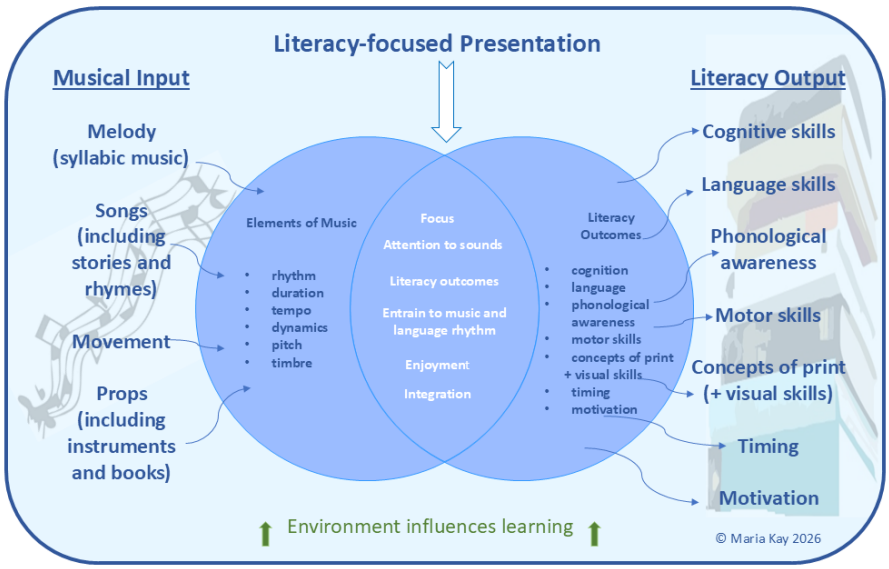
- Engagement – the activities held the participants' attention
- Attention drawn to sound and specifically language sounds
- Focus upon literacy outcomes
- Entrainment to a beat – tapping performance is linked to reading attainment (Tierney & Kraus 2013b; Tierney et al., 2021) as children must be able to process the temporal structure of sound in order to be aware of the sound structures in language (PA).
- Motivation and enjoyment – participants engaged readily with the activities and demonstrated enjoyment
- Integration – the activities promoted the ability of the participants to integrate their knowledge and skills and apply them to tasks set.

It is possible that other factors were also contributory to the popularity of activities, such as relevance to the children (boarding a bus in this example), ease of performance (the children had only to listen for their name and find a place on the "bus".

## Discussion

The data collected enabled the formulation of a LtM model (Figure 7 from Kay, 2026a) to explain how musical activities can be integrated with literacy learning to promote foundational literacy outcomes.

**Figure 7**
*Literacy-Through-Music Model (Kay, 2026a)*



Whilst the inputs are musical—the nature of which have been described in the intervention method—the outputs are foundational literacy outcomes. The underlying mechanisms that facilitated the successful promotion of these outcomes are central to the model.

The engaging nature of the activities and their suitability for children of this age, in terms of content and ease of participation facilitated the children's willingness to undertake the activities and focused their attention on the activities, the presenter and their performance. Children's attention was also focused upon discrimination of both musical and language sounds; the ability to discriminate language sounds being pivotal to literacy success.

The findings highlighted the importance of gesture in a child's early language development, both in terms of the child's response to the presenter's gestures before vocal communication and that of the child's response to questions posted by the presenter. It was clear to see that action songs and rhymes have an important role to play in helping to develop vocabulary and in the practising of language production, where gestures complement and support spoken language.

The integrative nature of both the activities and the skills being promoted allowed learning to be holistic and natural. Children were not learning isolated skills, but skills embedded within the performance of the tasks. The integration of music and literacy within activities obviated the need for skill transfer from one domain to another. For example, prosody in a musical phrase was reflected in the accompanying song lyrics, or the descending musical pitch at the end of verse was reflected in the falling intonation of the voice; thus, promoting the ability to add expression to language production. Similarly, the alignment of musical tones with language syllables naturally supported the children's ability to syllabify.

The repetitive nature of many of the songs and rhymes helped children to practise and refine their co-ordination skills and fostered the ability to entrain to a beat, through the performance of gross motor skills, such as stamping along in time to a song or fine motor skills when controlling and tapping claves. Fostering the ability to physically keep time aligned with vocal production strengthens auditory memory and supports language fluency which subsequently influences reading and writing fluency (Zanto et al., 2024). Fine motor skills were also promoted through the use of finger puppets. When children were required to perform together and at the same time, this further reinforced skill building as this synchronisation unconsciously invokes dynamic coupling, where the brain naturally mirrors the rhythms and movements of others; this further reinforced the ability to entrain to a beat and to consistently keep time.

The enjoyable nature of the activities allowed children to behave naturally, and their behaviour illustrated how children learn when they are relaxed and happy as they requested that activities be repeated and also, they repeated melodies and lyrics to themselves, thereby reinforcing any learning that was taking place.

The Literacy-through-Music Model illustrates how the focus of music-based activities can be upon literacy outcomes. It is incumbent upon the presenter or teacher to select and present a range of activities through which children can assimilate.

## Limitations

Limitations of the study were the duration of the study and the small sample size. A longitudinal study would be able to assess the longer-term benefits of implementation of a LtM program.

## Recommendations

The findings of this study suggest that using integrated music and literacy activities as a vehicle through which to deliver foundational literacy outcomes would be propitious. Kraus (2019) suggested that two years is the optimum time for musical activities to effectively make changes in the brain, in which case, preschool settings would benefit from implementing a LtM program for children aged three to five years of age. Furthermore, where children have not had the advantage of such a program then provision to bolster any gap should be made and LtM activities can be used in formal as well as informal school settings.

Many activities cover a broad spectrum of competencies simultaneously, but the presenter must be aware of the ones that are deliberately being promoted and note where this is happening and where it is not, such that revisions can be made in any program. The main predictors of literacy success should be specifically promoted—listening and focusing upon sounds, language comprehension and production, phonological awareness and the ability to play with the sound segments within words. The presenter must be aware of the skills to be promoted.

The sessions should be informal, and activities should be enjoyable and engaging; the teacher should be observing and promoting foundational literacy behaviours, not testing them. These features are present in the Finnish curriculum where music is an integral part of a school day from early years through primary. The Finnish Curriculum states that "Children have a right to learn through play and with joy" (Finnish National Board of Education, 2010). Assimilation of these features into English curricula is recommended.

## Conclusion

A unique contribution is made by this research into the benefits to the development of foundational literacy competencies through an integrated music and literacy program. It was able to address a gap in the literature of qualitative studies in early literacy learning through integrated music and literacy activities and it also bridged a gap between educational research and educational practice as the intervention was conducted as any other activity session in a normal day in a preschool setting.

Underlying the efficacy of the LtM program devised for this study is that the activities were enjoyable, engaging and motivational; they aroused interest and subsequent willingness to participate. Children learn best when they are interested (Harackiewicz et al., 2016). The children were relaxed and, in an environment, conducive to learning.

The differences between a music program and a LtM program lie predominantly in the focus upon literacy outcomes and the ability of the presenter to draw children's attention to sounds in spoken language.

There are multifarious benefits of music participation and adding these to literacy learning serves to augment the value of a LtM program and specifically the competencies identified in Figure 6. Their relationship is illustrated in Figure 7. The integration provided by a LtM

program engages both hemispheres of the brain simultaneously, enriching learning capacity and stimulating neuronal activity. The multi-sensory nature of the experiences also promotes embodied learning in a rich learning environment which is ideally suited to children of this age.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that no AI or AI-assisted technologies have been used to generate, refine or correct the content in the manuscript. The ideas, design, procedures, findings, analyses and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Berninger, V. W., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., Hawkins, J. M., & Graham, S. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology, 94*(2), 291–304. https://doi.org/10.1037/0022-0663.94.2.291

Berninger, V. W., & Winn, W. D. (2006). Implications of Advancements in Brain Research and Technology for Writing Development, Writing Instruction, and Educational Evolution. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (p. 96–114). The Guilford Press.

Bolduc, J., & Lefebvre, P. (2012). Using nursery rhymes to foster phonological and musical processing skills in kindergarteners. *Creative Education, 3*(4), 495–502. https://doi.org/10.4236/ce.2012.34075

Bostelman, T. J. (2008). *The effects of rhyme and music on the acquisition of early phonological and phonemic awareness skills* [Master's thesis, Defiance College]. OhioLINK Electronic Theses and Dissertations. https://etd.ohiolink.edu/acprod/odb_etd/ws/send_file/send?accession=def1281545878&disposition=inline

Cabrera, K. R., Gomez, R., & Franlin, J. (2024). *Early Literacy Matters: Economic impact and regional disparities in England*. Pro Bono Economics.

Carr, K. W., White-Schwoch, T., Tierney, A. T., Strait, D. L., & Kraus, N. (2014). Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proceedings of the National Academy of Sciences, 111*(40), 14559–14564. https://doi.org/10.1073/pnas.1406219111

Collins, A., Dwyer, R., & Date, A. (2020). *Music education: A sound investment*. A report commissioned by Alberts, The Tony Foundation.

Corriveau, K. H., Goswami, U., & Thomson, J. M. (2010). Auditory processing and early literacy skills in preschool and kindergarten population. *Journal of Learning Disabilities*, *43*(4), 369–382. https://doi.org/10.1177/0022219410369071

Dice, J. L., & Schwanenflugel, P. (2012). A structural model of the effects of preschool attention on kindergarten literacy. *Reading and Writing, 25*, 2205–2222. https://doi.org/10.1007/s11145-011-9354-3

Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly,* 56(S1), S25–S44. https://doi.org/10.1002/rrq.411

Finnish National Board of Education. (2010). *National core curriculum for pre-primary education*. https://www.oph.fi/en/statistics-and-publications/publications/new-national-core-curriculum-basic-education-focus-school

Fiveash, A., Ferreri, L., Bouwer, F. L., Kösem, A., Moghimi, S., Ravignani, A., Keller, P. E. & Tillmann, B. (2023). Can rhythm-mediated reward boost learning, memory and social connection? Perspectives for future research. *Neuroscience & Behavioral Reviews, 149.* https://doi.org/10.1016/j.neubiorev.2023.105153

Frasher, K. D. (2014). Music and literacy: Strategies using comprehension connections by Tanny McGregor. *General Music Today, 27*(3), 6–9. https://doi.org/10.1177/1048371314520968

Goswami, U. (2022). Language acquisition and speech rhythm patterns: an auditory neuroscience perspective. *Royal Society Open Science, 9*(7). https://10.1098/rsos.211855

Gough, P., & Tunmer, W. (1986). Decoding, reading and reading disability. *Remedial and Special Education, 7*, 6–10. https://doi.org/10.1177/0741932586007001

Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest matters: The importance of promoting interest in education. *Policy Insights from the Behavioral and Brain Sciences, 3*(2), 220–227. https://doi.org/10.1177/2372732216655542

Justice, L. M., & Kaderavek, J. N. (2004). Embedded-explicit emergent literacy intervention I: Background and description of approach. *Language, Speech and Hearing Services in Schools*, *35*, 201–211. https://doi.org/10.1044/0161-1461(2004/020)

Kay, M. (2026a). An Exploration of the Use of Musical Activities to Promote Foundational Literacy Competencies for Children Aged 3-5 Years. *The Reading Teacher, 80*(1). https://doi.org:10.1002/trtr.70068

Kay, M. (2026b). *My Cat Ben – Singalong.* Bryant and Kay Publishing.

Kraus, N. (2019). *Music and the Mind with Dr Nina Kraus.* [Video] YouTube. https://youtu.be/43I_PHyb_K8?si=TJe-fVeAosm9kaCl

Macrine, S. L., & Fugate, J. M. B. (2021). Translating embodied cognition for embodied learning in the classroom. *Frontiers in Education, Neuroscience, Learning and Educational Psychology.* https://doi.org/10.3389/feduc.2021.712626

Macrine, S. L., & Fugate, M. B. (Eds.). (2022). *Movement Matters: How embodied cognition informs teaching and learning.* MIT Press.

Martin, J., Bowl, M., & Banks, G. (Eds.). (2024). *Mapping the field: 75 Years of educational review, Volume 2*. Routledge.

Meyer, G. F., Spray, A., Fairlie, J. E., & Uomini, N. T. (2014). Inferring common cognitive mechanisms from brain blood-flow lateralization data: a new methodology for fTCD analysis. *Frontiers in Psychology, 5*(552). https://doi.org/10.3389/fpsyg.2014.00552

McWeeny, S., Choi, S., Choe, J., LaTourrette, A., Roberts, M. Y., & Norton, E. S. (2022). Rapid automatized naming (RAN) as a kindergarten predictor of future reading in English: A systematic review and meta-analysis. *Reading Research Quarterly, 57*(4), 1–20. https://doi.org/10.1002/rrq.467

Milne, N., Cacciotti, K., Davies, K., & Orr, R. (2018). The relationship between motor proficiency and reading ability in Year 1 children: a cross-sectional study. *BMC Pediatrics, 18*, 294. https://doi.org/10.1186/s12887-018-1262-0

Mirhosseini, S.-A. (Ed.) (2017). Introduction: Qualitative Research in language and Literacy Education in *Reflections on Qualitative Research in Language and Literacy Education* (pp. 1–13). Springer.

NLT (National Literacy Trust). (2026). *Annual Literacy Survey*. https://literacytrust.org.uk/research-services/annual-literacy-survey/

Novack, M., & Goldin-Meadow, S. (2015). *Learning from gesture: How our hands change our minds*. *Educational Psychology Review, 27*(3), 405–412. https://doi.org/10.1007/s10648-015-9325-3

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing, 30*(6), 653–661. https://doi.org/10.1097/AUD.0b013e3181b412e9

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology, Hypothesis and Theory Article, 2*(142). https://doi.org/10.3389/fpsyg.2011.00142

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research, 308,* 98–108. http://doi.org/10.1016/j.heares.2013.08.011

Polyanskaya, L., Samuel, A. G., & Ordin, M. (2019). Speech Rhythm Convergence as a Social Coalition Signal. *Evolutionary Psychology*, *17*(3). https://10.1177/1474704919879335

Putkinen, V., Tervaniemi, M., & Huotilainen, M. (2012). Informal musical activities are linked to auditory discrimination and attention in 2–3-year-old children - an event-related potential study. *European Journal of Neuroscience*, *37*, 654–661. https://doi.org/10.1111/ejn.12049

Rautenberg, I. (2015). The effects of musical training on the decoding skills of German-speaking primary school children. *Journal of Research in Reading, 38*(1), 1–17. https://doi.org/10.1111/jrir.12010

Register, D. (2001). The effects of an early intervention music curriculum on pre-reading/writing. *Journal of Music Therapy, 38*(3), 239–248. https://doi.org/10.1093/jmt/38.3.239

Rohde, L. (2015). The comprehensive emergent literacy model: Early literacy in context. *SAGE Open*, *5*(1). https://doi.org/10.1177/2158244015577664

Roulstone, S., Law, J., Rush, R., Clegg, J., & Peters, T. (2011). The role of language in children's early educational outcomes. *Research Brief. DFE-RB 134*, https://doi.org/10.1037/e603032011-001

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilites: Evidence, Theory, and Practice. In S. Neuman & D. Dickinson (Eds.). *Handbook for research in early literacy,* 97–110. Guilford Press.

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologica, 33*(8), 1047–1055. https://doi.org/10.1016/0028-3932(95)00045-5

Sedita, J. (2019). *The Strands That Are Woven Into Skilled Writing*. https://keystoliteracy.com/wp-content/uploads/2020/02/The-Strands-That-Are-Woven-Into-Skilled-WritingV2.pdf

SERA. (2005). *Scottish Educational Research Association Ethical Guidelines for Educational Research*. SERA.

Slater, J., Strait, D. L., Skoe, E., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research*, *291*, 244–252. https://doi.org/10.1016/j.bbr.2015.05.026

Slater, J., Tierney, A., & Kraus, N. (2013). At-risk elementary school children with one year of classroom music instruction are better at keeping a beat. *PLoS ONE*, *8*(10), e77250. https://doi.org/10.1371/journal.pone.0077250

Sosu, E., & Ellis, S. (2014). *Closing the Attainment Gap in Scottish Education*. Joseph Rowntree Foundation.

Standley, J. M., & Hughes, J. E. (1997). Evaluation of an early intervention music curriculum for enhancing pre-reading/writing skills. *Music Therapy Perspectives*, *15*, 79–85. https://doi.org/10.1093/mtp/15.2.79

Stein, J. (2022). The visual basis of reading and reading difficulties. *Frontiers in Neuroscience, 16,* 1004027. https://doi.org/10.3389/fnins.2022.1004027

Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children. The role of language comprehension and working memory skills. *Reading and Writing, 4*, 245–256. https://doi.org/10.1007/BF01027150

Tierney, A., & Kraus, N. (2013a). Music Training for the Development of reading skills. In M. Merzenich, M. Nahum and T. Van Vleet (Eds.). *Progress in Brain Research*, 209–241. Academic Press.

Tierney, A. T., & Kraus, N. (2013b). The ability to tap to a beat relates to cognitive, linguistic and perceptual skills. *Brain and Language*, *124*, 225–231. https://doi.org/10.1016/j.bandl.2012.12.014

Tierney, A. T., & Kraus, N. (2014). Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Frontiers in Human Neuroscience, 8*, 949. https://doi.org/10.3389/fnhum.2014.00949

Tierney, A. T., Gomez, J. C., Fedele, O., & Kirkham, N. Z. (2021). Reading ability in children relates to rhythm perception across modalities. *Journal of Experimental Child Psychology, 210.* https:// 10.1016/j.jecp.2021.105196

Toyka, K. V., & Freund, H. J. (2007). Music, motor control and the brain. *Brain, 129*(10), 2794–2798. https://doi.org/10.1093/brain/awl231

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, & National Assessment of Educational Progress. (2025). *The nation's report card: Grade 12 Reading.* https://www.nationsreportcard.gov/reports/reading/2024/g12/

Verney, J. P. (2013). *Rhythmic perception and entrainment in 5-year-old children: An exploration of the relationship between temporal accuracy at four isochronous rates and its impact on phonological awareness and reading development* [Doctoral dissertation, University of Cambridge]. Apollo – University of Cambridge Repository. https://doi.org/10.17863/CAM.16494

Winkler, I., Haden, G., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn Infants Detect the beat in music. *Proceedings of the National Academy of Sciences, 106*(7), 2468–2471. https://doi.org/10.1073/pnas.0809035106

Zanto, T. P., Giannakopoulou, A., Gallen, C. L., Ostrand, A. E., Younger, J. W., Anguera-Singla, R., Anguera, J. A., & Gazzaley, A. (2024). Digital rhythm training improves reading fluency in children. *Developmental Science, 27*(3). https://doi:10.1111/desc.13473

Zhakim, A. (2024). Pre-school children's learning and development of literacy in the Kazakh home learning environment. *Journal of Education and E-Learning Research, 11*(1), 142–154. https://doi.org/10.20448/jeelr.v11i1.5398

**Contact email:** a.maria.kay@btinternet.com

# From "Curriculum as Product" to "Curriculum as Process": An Autoethnographic Critical Review of English Language Curriculum Development in Hong Kong

Shuk Ling Cheng, The Hong Kong Polytechnic University, Hong Kong
Wenyun Jia, City University of Hong Kong, Hong Kong

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

This autoethnographic paper critically examines the development of English language curriculum in Hong Kong secondary schools by juxtaposing the author's educational experience with the principles outlined in the English Language Curriculum and Assessment Guide (2021). Drawing on curriculum theory, this paper analyzes the transition from traditional "curriculum as product" and "curriculum as knowledge" models, characterized by examination-oriented instruction, discrete skill assessment, and behaviorist pedagogy, towards the process-centered and learner-centered approaches advocated in the current curriculum framework. This paper identifies a persistent gap between official curriculum intentions and classroom practices, noting that product-oriented methodologies continuously dominate despite the fact that post-reform education policies in Hong Kong nowadays emphasize learner autonomy, communicative competence, and formative assessment. Pedagogical implications include the need for systemic support in teacher training, assessment reform, and resources allocation to facilitate genuine curriculum transformation. This paper argues for a dialectical approach that integrates both product and process orientations, enabling meaningful language learning while addressing contextual constraints in Hong Kong's educational landscape.

*Keywords:* autoethnography, curriculum development, English language education, Hong Kong secondary schools, pedagogical implications

iafor
The International Academic Forum
www.iafor.org

## Introduction

Through examining my personal experience studying at a government-funded English Medium of Instruction (EMI) secondary school from 2008 to 2014, this review paper offers a critical account of analyzing how the curriculum approach I encountered in my school years compares with the contemporary framework in English Language Curriculum and Assessment Guide (2021). Curriculum theory—specifically examining the conceptualizations of curriculum as knowledge, product, process, and praxis—is applied for considering the implications of devising English language education in Hong Kong.

My local school had paid a strong focus on promoting an English-rich learning environment through organizing English Speech Day, English Reading Week, Movie Society and Debate Team. Nonetheless, from the standpoint of curriculum development, my school's approach perennially perpetuated the traditional examination-oriented teaching system that emphasizes considerably discrete language skills (predominantly grammar) and measurable outcomes. The pedagogical framework combined elements of structural-situational methodology with task-based learning, though the former appeared to dominate classroom practice.

Under this curriculum structure, two main components, namely grammar and vocabulary building occupied a central position. First, acquisition of grammar was usually consolidated through regular quizzes designed to establish what teachers termed our grammatical foundation. These assessments followed rudimentary patterns: presentation of grammatical rules (including prescriptive and descriptive grammar), controlled practice exercises and testing of specific structures in isolation. Explicit teaching of grammar that functions within genre writing is particularly popular in Hong Kong teaching (Halliday, 2003). Second, the building of vocabulary was addressed primarily through dictation and vocabulary quizzes. We were required to memorize predetermined lists of phrases and expressions deemed useful for DSE English writing, which had a set focus on particular topics, such as social issues, popular culture and sports communication. At my school, there were frequent examinations on developing reading and writing skills that trained us to respond to specific question formats and produce writing articles specially complying with DSE examination rubrics.

In addition to examinations and quizzes, the School-Based Assessment (SBA) component represents a kind of language learning that involves the acquisition of linguistic and communication skills. We were required to write reviews of books and deliver presentations to the class. Following each presentation, the English teacher posed spontaneous questions requiring students to provide immediate responses. While these activities ostensibly engaged multiple skills and promoted higher-order thinking skills, the implementation was very likely constrained by assessment criteria and time pressures. More often than not, SBA (although exposing students in a relatively low-stakes context) in English classrooms seems to amount to nothing more than regular summative testing, grading and record keeping of marks to prepare students for high-stakes examinations, given that SBA accounted for 15% of the entire English DSE grade at my time. The tasks in general were formative and similar administered subsequently as part of the public assessment process to generate marks for standardizing the learning progress.

## Theoretical Underpinnings of English Language Curriculum and Assessment Guide (2021) and Its Relevance

It is worth noting that the 2021 English Language Curriculum and Assessment Guide emphasizes flexibility, learner autonomy and holistic development. The guide principle comprises three interlocking components: subject knowledge and skills; generic skills; positive values and attitudes. This tripartite structure signals a departure from purely linguistic competence toward broader educational outcomes that guarantees the promotion of interpersonal development, knowledge and experiential engagement in English language acquisition.

Significantly, the 2021 guide advocates task-based learning as a central pedagogical approach, with a focus on "learning to communicate through purposeful interaction" and engagement in "processes that require them to exercise critical thinking and creativity" (CDC-HKEAA, 2021, p. 30). Such an approach reveals a constructivist orientation, in which students construct their own understanding and knowledge of the world through making meaningful connections and critically reflecting on experiences. This outcome is further substantiated when language learning is described as "a dynamic, interactive process in which students play an active role in using language to make sense of the world" (CDC-HKEAA, 2021, p. 29). It is clear that the guide reflects a more formative and learner-centered vision in assessment for learning and assessment as learning, that empowers students to "reflect on and monitor their progress" (CDC-HKEAA, 2021, p. 62).

To explore the relationship between my educational experience and the 2021 framework, it is helpful to employ the typology of curriculum approaches articulated by Smith (1996) and Finney (2002), which distinguishes curriculum as knowledge, product, process and praxis. Curriculum as knowledge is the more traditional and passive ways of transmission of established content and facts (Posner, 2004), with authorities (e.g. teachers) delivering predetermined learning items to recipients (e.g. students). Curriculum as product views curriculum development as systematic specification of educational purposes. It focuses on the learning experiences to achieve them, actual interventions and evaluation methods (e.g. providing feedback and changes) to determine whether purposes have been attained. Curriculum as process shifts from predetermined outcomes to the learning journey itself (Stenhouse, 1975). This approach emphasizes what transpires in classrooms, encourages experimentation and critique, and recognizes that curriculum is unique for each teaching context: learners occupy a voice in learning while teachers judge the effectiveness of teaching and adapt (Smith, 1996). Finally, curriculum as praxis extends the process model, which deals with practical deliberation and differentiated curriculum. Grundy (1987) states that curriculum is "constituted through an active process in which planning, acting and evaluating are all reciprocally related and integrated". This means that there is always a collective need to tailor made teaching environments and practices that create appropriately different learning experiences for students, who can then feel safe and flexible to engage their abilities to the greatest extent possible.

My educational experience mostly reflected the concepts of "curriculum as product" and "curriculum as knowledge". The utilization of grammar quizzes, dictation exercises and examination reveal an objectives-driven model centered on measurable linguistic outcomes. The regular-basis testing of grammatical structures in isolation, coupled with memorization of predetermined phrases for examination writing, exemplifies Tyler's rational curriculum model (1949): specific language items were identified as objectives; teaching activities were designed

to facilitate their acquisition; and assessment measured whether students had mastered these skills necessary for examination. All these point to the behaviourist pedagogy of my school's English teaching, where the focus is on "intended products" and "teaching to help acquisition of products" (Cunningham et al., 2007). In addition, the repetition of grammar-translation methodology in structural practice indicates a generic view of language learning as accumulation of linguistic knowledge rather than development of communicative competence. Although my school's SBA seemed to incorporate some formative elements, such as teacher's feedback and opportunities for oral interaction, its prime function remained summative assessment for public examination. The emphasis on training to produce specific response formats constrained the potential for genuine formative assessment supporting genuine communication skills and interpersonal interaction.

In contrast, the 2021 guide positions itself as "curriculum as process" and subtly touches upon "curriculum as praxis". According to the guide, teachers and schools should be given flexibility and ownership in curriculum planning, which suggests that effective curriculum emerges through situated practice rather than generic prescription applied to all. The guide also pays attention to values, attitudes and critical engagement at the level of praxis, where curriculum should be more activity- and metacognitively centred. The encouragement to develop students' "critical thinking skills" and ability to "evaluate critical views and attitudes" around their living (CDC-HKEAA, 2021, p. 38) also suggests concern for learner empowerment beyond functional linguistic competence.

In terms of the roles at school, I as a student was positioned primarily as recipients of transmitted knowledge and performers of prescribed tasks. The grammar-translation approach and examination focus afforded limited opportunities for genuine choice, creativity, or critical engagement. We learned predetermined grammatical structures, memorized selected vocabulary, and practiced producing texts conforming to examination rubrics. This reflects the behaviourist metaphor of learning as acquisition, reproduction, imitation, modelling among students who are dependent on teachers (Cunningham et al., 2007). But the 2021 guide envisions substantially greater student agency in autonomous learning. Instead of being a passive learner, students are encouraged to develop independence and self-directed learning skills in order to set goals and make informed choices as to what, when and how they want to learn. The guide advocates engaging students in self-assessment and self-reflection, which can help them plan, monitor and evaluate the strategies adopted and their own learning effectiveness. Such a vision is fortified by the curriculum as process which recognizes learners as active meaning-makers and signals the commitment to cultivating self-motivated learners rather than passive recipients of instruction.

Finally, the pedagogy I experienced revealed a blend of structural-situational methodology and emerging task-based approaches, though implementation favored the former. The Presentation-Practice-Production (PPP) sequence was evident in my grammar lessons. The English teachers presented rules while we as students completed controlled exercises and were tested on specific structures. This reflected the audio-lingual tradition's behaviourist theory upholding that language learning is a habit formation through stimulus-response conditioning. Meanwhile, the dictation exercises similarly reveal behaviourist influences, which centre on the practices of mimicking and receiving instruction from the human role models (Cooter & Reutzel, 2004). In this way, vocabulary acquisition is grasped through memorization of decontextualized items and consistent repetition. This approach contradicts the 2021 guide suggesting that vocabulary is best introduced in context using reading, listening or multimodal texts, and practised through tasks and language games. It is claimed that students should be consistently provided with

purposeful contexts and engaged in processes that require them to exercise critical thinking and creativity reflects understanding of language learning as active knowledge construction, rather than a process of habit formation in which English patterns were presented for memorization. This student-centred instruction and promotion of self-directed learning strategies indicate alignment with a more diversifying learning principle that values learner autonomy, intrinsic motivation and individual development.

The framework explicitly critiques approaches like those I experienced, noting that "formal explanation of grammatical rules in isolation and the use of decontextualized and mechanical drills are not always useful in helping students develop communicative competence" (CDC-HKEAA, 2021, p. 32). Instead, it advocates teaching grammar "in context" within meaningful tasks, positioning grammatical knowledge as means to communicative ends rather than ends in themselves.

## Discussion and Conclusion

From my personal account, it is obvious that my educational experience in a mainstream government school suggests substantial gaps between official curriculum vision and classroom reality. While the 2021 guide represents an ideal vision for an interactive teaching pedagogy rather than radical departure from earlier curriculum documents, the predominantly product-oriented, examination-focused approach I encountered indicates that progressive curriculum rhetoric often fails to transform classroom practice. My education experience had unravelled a type of knowledge-transmission pedagogy that entailed predetermined linguistic objectives, examination preparation and measurable outcomes. In contrast, the 2021 guide places a great emphasis on ideally process-oriented approaches which uphold meaningful communication, learner agency, flexible organization and formative assessment. Such an implementation gap likely reflects multiple factors: examination pressure; teacher preparation and beliefs about effective language teaching; resource limitations; class sizes; and the path-dependent nature of educational institutions. As Finney (2002) observes, teachers inevitably interpret and adapt curriculum frameworks through their existing pedagogical commitments and contextual constraints.

In light of this, curriculum development may require supportive implementation structures. Progressive curriculum visions necessitate aligned teacher education, assessment systems, resource provision and school cultures. Without these supports, such a developmental plan risks remaining rhetorical exercises. Genuine curriculum transformation requires examination reform that authentically assesses the communicative competence, critical thinking and creativity the guide intends to espouse. Additionally, teacher agency and professional judgment play a pivotal role in effective curriculum implementation. How to localize the context and make it adaptive or tailor-made to students requires investment in teacher education and creation of school cultures supporting pedagogical experimentation. While it is understandable that teachers have been juggled with many tasks other than teaching and designing curriculum, there needs to be more resources (e.g. manpower) allocated to avoid teachers' work overload and increased pressure. Last but not least, in lieu of viewing product and process orientations as mutually exclusive, effective curriculum development might consider embracing their dialectical relationship. This may facilitate meaningful learning processes that can lead to valuable knowledge and skill products, while acknowledging students' need for linguistic tools to participate in valued communicative practices.

Reflecting on my own language learning journey, I recognize both the limitations of the product-oriented approach I experienced and the foundation it provided. The grammatical knowledge and examination skills I acquired have served me well, though I have often needed to develop communicative fluency, critical thinking skills and learner autonomy through advanced study and professional experience.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Generative AI was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The ideas, design, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Cunningham, B., Dawes, G., & Bennett, B. (2007). *The handbook of work based learning*. Routledge.

Curriculum Development Council & Hong Kong Examinations and Assessment Authority. (2021). *English Language Curriculum and Assessment Guide (Secondary 4-6)*. Education Bureau, HKSAR.

Finney, D. (2002). The ELT curriculum: A flexible model for a changing world. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 69–79). Cambridge University Press.

Grundy, S. (1987). *Curriculum: Product or praxis?* Falmer Press.

Halliday, M. A. K. (2003). Written language, standard language, global language. *World Englishes*, *22*(4), 405–418.

Posner, G. J. (2004). *Analyzing the curriculum* (3rd ed.). McGraw-Hill.

Reutzel, D. R., & Cooter, R. B. (2004). *Teaching Children to Read: Putting the Pieces Together* (4th ed.). Merrill/Prentice-Hall Publishing Company.

Smith, M. K. (1996/2000). *Curriculum theory and practice*. The Encyclopaedia of Informal Education.

Stenhouse, L. (1975). *An introduction to curriculum research and development*. Heinemann.

Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. University of Chicago Press.

# Spaces That Follow Individual Preference Can Improve Learning Performance

Walaa Aljehani, Lancaster University, United Kingdom
Ana Rute Costa, Lancaster University, United Kingdom
Michelle To, Lancaster University, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

The design of learning environments has traditionally focused on standardised spaces and layouts. However, recent research suggests that individual spatial preferences may affect learning outcomes. This paper discusses the relationship between individual spatial preferences and learning performance in learning environments. In a virtual reality setting, participants (N = 60) completed a memory task in different virtual learning environments. A mixed-methods approach was employed to analyse the interaction between individual spatial preferences and memory performance across these virtual learning environments. Learning performance was assessed using recall tests to measure memory retention and a single open-ended question to explore spatial preferences. The learning environment design conditions comprised 6 variations: 2 lighting temperatures (3000K, 6000K), 2 colour palettes (warm, cold), and 2 room proportions (large, small). The six rooms all shared the same design, with one parameter modified in each condition. The study shows that designing spaces based on individual preferences can enhance learning performance. Findings reveal a noticeable interaction between individual spatial preferences and memory performance. Participants whose spatial preferences aligned with the spatial configuration had retention scores 79.2% higher than those in non-preferred environments. Moreover, this preference-performance relationship remained consistent across all conditions: 85.7% in the room condition, 77.8% in the light condition, and 75% in the colour condition. These findings deepen our theoretical understanding of person-environment fit in learning environments, showing that adaptable spatial designs can improve learning outcomes. This approach promotes evidence-based, user-centred design strategies that optimise educational spaces for diverse cognitive needs, shifting from one-size-fits-all solutions to more flexible, customizable and adaptive learning environments.

*Keywords:* learning environment, individual preference, spatial design, academic performance

iafor
The International Academic Forum
www.iafor.org

## Introduction

Numerous studies have proposed evidence-based designs for educational spaces to enhance learning outcomes (Barrett et al., 2015; Byers et al., 2018; Kariippanon et al., 2018). Still, historically, there has been a focus on prioritising functional standardisation over accommodating individual variations, resulting in universal design in learning environments calibrated for an assumed average user rather than catering to the diverse cognitive and perceptual profiles of actual learners (Rose & Strangman, 2007).

However, a substantial body of research in environmental psychology and educational neuroscience indicates that people respond to spatial stimuli in markedly different ways (Aagaard, 2018; Costa, 2025; Fraser & Fisher, 1983; Greeno, 1994; Kozhevnikov et al., 2014). Therefore, it is pedagogically significant to acknowledge and consider the variability in individual differences and preferences within learning environments to effectively achieve specified learning objectives (Cronbach & Snow, 1981).

Research and educational institutions increasingly improve learning environments that can be flexibly reconfigured in terms of furniture arrangement, lighting control, and aesthetic character (Osborne, 2016; Temple, 2008). A real-world example of this concept is Vittra School Telefonplan in Stockholm, where Rosan Bosch Studio replaced conventional classrooms with a versatile "learning landscape" consisting of activity-based zones that students can freely navigate based on their personal needs (Rosan Bosch Studio, n.d.). However, these adaptations are rarely evaluated in terms of individualised preferences and their effect on learning outcomes. The dominant discourse around flexible learning design focuses on collaborative versus independent work modes, rather than on cognitive and spatial personalisation.

This study evaluates whether aligning spatial conditions with participants' self-reported spatial preferences enhances performance on standardised cognitive tasks (memory tasks). Using a Virtual Reality (VR) setting, the relationship between spatial preference and performance was investigated across three variables: lighting, colour palette, and room proportion. The study involved 60 university students who were exposed to two distinct experimental conditions within a university study room setting, each participant experiencing two different designs for informal academic study environments.

## Literature Review

### Person–Environment Fit in Educational Settings

Person–Environment fit (P-E fit) theory provides a foundational framework for understanding why alignment of spatial preferences may influence academic performance (De Cooman & Vleugels, 2022). The theory has a long history, traceable to early work on congruence in vocational counselling, and has since been developed and expanded by numerous scholars to explain how the interaction between personal characteristics and the surrounding environment influences individual outcomes (Aagaard, 2018a; Greeno, 1994; Kozhevnikov et al., 2014). At its core, a fit hypothesis suggests that compatibility between a person and their environment leads to better individual outcomes compared to when they are mismatched (Zhang et al., 2025).

In educational contexts, P-E fit theory states that a person's satisfaction and success depend on the degree of congruence between the individual's characteristics and the environment's characteristics. Greater alignment leads to higher satisfaction, success, and retention. Specifically, a good fit between personal factors—such as learning preferences and abilities—and environmental features results in better study performance and higher satisfaction (Fraser & Fisher, 1983).

### Environmental Psychology and Learning

Research in environmental psychology shows that physical surroundings have a measurable impact on mood, cognition, and behaviour (Russell & Ward, 1982). In education, studies have linked environmental factors like natural light, sound quality, air quality, and space density to outcomes such as reading comprehension, math skills, and attendance (Barrett et al., 2015; Choi et al., 2014). Despite these findings, much research treats the physical environment as a uniform factor, assuming that better lighting benefits all students equally. However, less attention has been paid to whether personal preferences affect how the environment influences performance.

### Cognitive Affordances in Learning Spaces

Numerous studies have shown that individuals actively interpret and shape their environments, leading to different cognitive affordances even within the same physical space. This variation is shaped by personal meaning, attention, and prior experience (Gibson, 2014; Kiverstein & Artese, 2024). Affordances in learning environments are not fixed properties of space but relational qualities that emerge from the interaction between an individual and their surroundings. Gibson (1979) established that affordances depend on the actor's perception, meaning the same physical space can offer fundamentally different cognitive opportunities to different individuals. Building on this, Jelić et al. (2016) propose that architectural affordances arise from the coupling between a person's body schema and the environment's action possibilities, arguing that perception, action, and emotion are dynamically intertwined, so the same space yields different affordance experiences depending on physical capacities, prior experiences, and emotional states. This is consistent with Bandura (1986) the concept of reciprocal determinism, whereby personal factors, behaviour, and environmental influences interact continuously, with individuals both shaping and being shaped by the environments they encounter.

## Methodology

### Participants

A total of 60 university students participated in the study (20 per environmental condition group). Participants were recruited through voluntary sign-up at Lancaster University. Inclusion criteria required the absence of self-reported visual, neurological, or psychological conditions that might confound environmental sensitivity measures. Ethical approval was obtained from the institution's research ethics committee prior to data collection, and all participants provided written informed consent.

## Research Design

The study employed a within-subjects design in which each participant was assessed under both a preferred and a non-preferred condition for their assigned environmental variable. Participants were randomly allocated to one of three environmental variable groups: lighting (n = 20), colour (n = 20), or room proportion (n = 20). Prior to experimental sessions, all participants engaged in semi-structured interviews and discussed their self-reported optimal conditions within their assigned variable.

Cognitive performance was assessed using a short-term memory recall test. The test used a diverse set of words, frequently employed in free recall studies. Qualitative data were also collected through a single open-ended question administered after each session, providing deeper insight into participants' experiences and perceptions of the learning environment. Participants were asked: How different is your learning experience in the two conditions? Which one do you prefer?

## Environmental Conditions

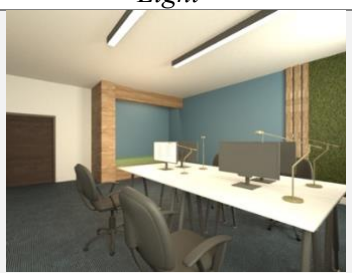
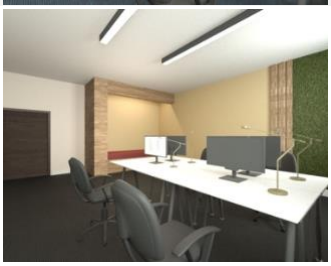
The six study rooms were designed based on previous literature reviews and current research to create an optimal learning environment. The review identifies interior elements that positively impact learning and memory. Lighting conditions were operationalised as two levels: warm light (3000K) and cool light (6000K). Colour conditions comprised two palette options: warm tones (yellow, red, orange, brown), and cool tones (blue, green, purple, navy). Room proportion conditions presented participants with spaces of a small room (3 × 4 m, 3 m ceiling height) and a large room (6 × 8 m, 5 m ceiling height). All other environmental factors (furniture, layout) were held constant across conditions. Table 1 describes the experiment's phases and two conditions in each phase. While Table 2 presents the image of the rendered perspective for each phase.

**Table 1**

*Describes the Experiment's Phases and Two Conditions in Each Phase*

| *Phase* | *Environmental variable* | *Condition* | *Specifications* |
|---|---|---|---|
| *Phase 1 Light* | Lighting colour temperature | A- Warm light | 3000K |
| | | B- Cold light | 6000K |
| *Phase 2 Colour* | Wall / interior colour palette | A- Cold colour | Blue, green, purple, navy |
| | | B- Warm colour | Red, yellow, orange, brown |
| *Phase 3 Room proportions* | Room dimensions & ceiling height | A- Small room | 3 × 4 m, 3 m ceiling height |
| | | B- Large room | 6 × 8 m, 5 m ceiling height |

**Table 2**
*Presents the Rendered Perspective Image for Each Phase*

| *Phase* | *Light* | *Colour* | *Room proportions* |
|---|---|---|---|
| *A* |  | | |
| *B* | |  | |

## Data Analysis

Data were analysed using a preference–performance alignment criterion: a participant was classified as showing positive alignment if their memory task score was higher in the preferred condition than in the non-preferred condition. The primary outcome measure was the proportion of participants demonstrating clear positive alignment within each group and across the full sample. Quantitative data were analysed using paired-sample t-tests to compare performance and subjective measures between the two conditions.

Qualitative data from post-session interviews were transcribed verbatim and analysed using Braun and Clarke (2006) a six-phase thematic analysis framework. Transcripts were first read repeatedly to familiarise with the data. Initial codes were then generated inductively using NVivo software, allowing themes to naturally emerge from participants' accounts instead of being driven by pre-existing theoretical categories. Codes were subsequently collated into candidate themes, which were reviewed and refined against both individual coded extracts and the dataset to ensure internal coherence and clear distinction between themes. This iterative process resulted in three final themes: Cognitive Comfort & Focus, Environmental Adaptation or Resistance, and Environmental Neutrality, each capturing a distinct pattern in participants' qualitative accounts of their environmental experience.

## Preference–Performance Alignment

Preference–performance alignment was assessed by comparing each participant's objective memory test performance across the two conditions with their stated environmental preference. A participant was coded as aligned (Yes) if their preferred condition matched the one in which they achieved a higher memory score, and not aligned (No) if their preferred condition matched the one in which they achieved a lower memory score.

Two categories of participants were excluded from the alignment rate calculation. First, those with identical memory scores in both conditions were labelled as Equal, because no meaningful performance difference could be identified. Second, participants who did not express any environmental preference were labelled as None, since alignment cannot be assessed without

a stated preference. These cases were considered uninformative regarding the preference–performance relationship and were excluded from the total count.

## Results

### Quantitative Analysis

#### *Memory Performance Across Environmental Conditions*

Paired-samples t-tests compared memory scores across conditions for each environmental variable. No significant differences were found among the three variables. Effect sizes were negligible to small across all three comparisons, consistent with the absence of significant results.

**Table 3**
*Paired-Sample t-Test Results: Memory Performance Mean and Standard Deviation*

| *Variable* | *Condition* | *M* | *SD* |
|---|---|---|---|
| *Lighting* | Cold light | 10.60 | 2.68 |
| | Warm light | 10.45 | 1.73 |
| *Colour* | Cold colour | 11.50 | 2.63 |
| | Warm colour | 11.65 | 2.23 |
| *Room proportion* | Large room | 11.15 | 3.03 |
| | Small room | 10.80 | 2.75 |

**Table 4**
*Paired-Samples t-Test Results for Memory Performance Across Conditions*

| *Variable* | *t* | *df* | *p* | *Mean diff.* | *SD diff.* | *Cohen's d* |
|---|---|---|---|---|---|---|
| ***Lighting*** | *0.231* | 19 | *.820* | 0.15 | 2.91 | *0.05* |
| ***Colour*** | *−0.318* | 19 | *.754* | −0.15 | 2.11 | *−0.07* |
| ***Room proportion*** | *0.733* | 19 | *.472* | 0.35 | 2.13 | *0.16* |

Whilst no condition produced a statistically superior memory outcome at the group level, examination of individual preference–performance alignment revealed a markedly different pattern.

#### *Preference–Performance Alignment*

Across all 60 participants, 38 demonstrated clear preference–performance alignment, yielding an overall alignment rate of 79.2%. Results varied by environmental variable, as summarised in Table 5.

**Table 5**
*Preference–Performance Alignment by Environmental Variable*

| *Environmental Variable* | *Total (n)* | *Aligned (Yes)* | *Not Aligned (No)* | *Excluded (Equal+None)* | *Denominator (Yes+No)* | *Alignment Rate* |
|---|---|---|---|---|---|---|
| *Lighting* | 20 | 14 | 4 | 2 | 18 | 77.8% |
| *Colour* | 20 | 12 | 4 | 4 | 16 | 75.0% |
| *Room Proportion* | 20 | 12 | 2 | 6 | 14 | 85.7% |
| *TOTAL* | 60 | 38 | 10 | 12 | 48 | 79.2% |

### *Lighting*

Of the 20 participants, 14 (70%) performed better under their preferred lighting condition. After excluding participants who recorded equal performance across conditions or expressed no lighting preference, the preference–performance alignment rate was 77.8% (14 of 18 participants).

### *Colour*

Eleven of 20 participants (55%) showed alignment between colour preference and memory performance. After excluding those with equal performance or no stated colour preference, the alignment rate rose to 75.0% (12 of 16 participants).

### *Room Proportion*

Twelve of 20 participants (60%) performed better in their preferred room-size condition, and room proportion also produced the highest rate of equal performance outcomes across the three variables (n = 5, 25%). After excluding equal-performance and no-preference cases, room proportion produced the strongest preference–performance alignment of the three variables, at 85.7%.

### *Inferential Analysis*

Chi-square goodness-of-fit tests were conducted to assess whether the observed rate of preference–performance alignment exceeded the 50% rate expected by chance. Alignment was significantly greater than chance for all three environmental variables. For lighting, 77.8% of participants performed better in their preferred condition, $\chi^2(1) = 5.556$, $p = .018$, $\varphi = 0.556$, indicating a medium effect. For colour, 75.0% of participants showed alignment, $\chi^2(1) = 4.000$, $p = .046$, $\varphi = 0.500$, again reflecting a medium effect. The strongest effect was observed for room proportion, where 85.7% of participants performed better in their preferred condition, $\chi^2(1) = 7.143$, $p = .008$, $\varphi = 0.714$, representing a large effect. Across all environmental variables combined, preference–performance alignment was observed in 79.2% of cases, significantly exceeding the chance level of 50%, $\chi^2(1) = 18.083$, $p < .001$, $\varphi = 0.614$, constituting a medium-to-large effect.

## Qualitative Analysis

The interview data analysis showed distinct differences in experiences between the conditions. Transcripts were examined using an inductive, reflexive thematic analysis method (Braun & Clarke, 2006). Through iterative clustering and refinement, three primary themes emerged. These themes include:

### *Theme 1: Cognitive Comfort & Focus*

This theme highlights the (28 of 48) 79.2% "Aligned" Groups who performed better in their preferred setting. Interviews show that when their preference was matched, it helped them feel comfortable, reduced distractions, and improved their concentration on the memory task.

### *Theme 2: Environmental Adaptation or Resistance*

The explanation of the (10 of 48) 20.8% "No Correlation" group illustrates why certain participants perform better in environments they have reported disliking. The interviews provide conflicting feedback on what they felt and how they performed the task, suggesting they may have distracted the effort to concentrate.

### *Theme 3: Environmental Neutrality*

An explanation for the (12 of 60) 20% the "Equal" group show that they do not pay attention to their surroundings, emphasise aesthetics or their environment when studying; instead, they concentrate mainly on the task itself.

Table 6 combines the quantitative alignment results with participants' qualitative comments, showing representative statements for each performance group (Aligned, Not Aligned, Equal). It also includes the relevant theme and a cognitive or environmental explanation for each pattern.

## Integration Results

To explain the reasons for and individual differences in quantitative performance, a thematic analysis of the semi-structured interviews was conducted. The qualitative results were mapped directly onto the three performance groups (Aligned, not aligned, and no preference/Equal performance) using a thematic matrix (Table 6). This approach shows that although performance metrics indicate what happened during the VR task, subjective experiences, such as cognitive comfort, compensatory arousal, and environmental neutrality, may explain why these spatial preferences affected learning outcomes.

**Table 6**
*Mixed-Methods Integration of Preference–Performance Alignment: Themes, Representative Quotes, and Explanations*

| PERFORMANCE GROUP | PRIMARY THEME | SUB-THEMES | REPRESENTATIVE PARTICIPANT QUOTES | COGNITIVE / ENVIRONMENTAL EXPLANATION |
|---|---|---|---|---|
| **Aligned**<br>Better performance in preferred space<br>79.2% of sample | **Cognitive comfort and focus** | · Preference-environment alignment<br>· Stress and anxiety reduction<br>· Preference-facilitated attention | *"I prefer the warm setting. In the cold light, I don't feel comfortable - there was something wrong for me. The warm light was better; I like it more."* (P-Light-03)<br>*"I prefer the cold colour. I was more comfortable with it. I don't like the warm colour palette; it distracts me."* (P-Colour-14)<br>*"I prefer the first one - the cold colour. I like the blue and green; I thought it calmed me down and helped me remember the words."* (P-Colour-17)<br>*"I feel I produced better results and felt more confident in the smaller room."* (P-Room-02) | Preference-performance alignment minimises spatial anxiety and reduces extraneous cognitive load. When the physical environment conforms to individual expectations, attentional resources are fully directed toward the memory task rather than environmental appraisal.<br>*person–environment fit theory; cognitive load theory (Kristof-Brown et al., 2005)* |
| **Not aligned**<br>Worse performance in preferred space<br>20.8% of sample | **Environmental adaptation or resistance** | · Arousal-mediated engagement<br>· Opposed attentional effort | *"I definitely felt more comfortable in the first setting with the cold light, but I was a lot more distracted and kept looking around more."* (P-Light-02)<br>*"I prefer the cold colour environment, but I think I performed better in the warm colour condition."* (P-Colour-09)<br>*"When I tried to recall the vocabulary, I didn't really notice the environment because I needed to focus on the task; otherwise, I couldn't remember."* (P-Room-13)<br>*"In the big room I simply had more to get distracted by in the environment."* (P-Room-14) | The preferred environments may function as mild stressors, triggering heightened physiological arousal that may reduce the task concentrate. |
| **Equal**<br>No performance difference across conditions | **Environmental neutrality** | · Task-driven attentional override<br>· Spatial indifference | *"I felt the warm light was more focused but found the cold light more comfortable."* (P-Light-01)<br>*"The cold light was clearer and easier for learning words, but the warm light provided a more relaxing environment."* (P-Light-06)<br>*"I didn't notice any difference. The first one felt lighter and brighter; the second had a lower light level, which made the task slightly more difficult."* (P-Colour-03)<br>*"I haven't noticed a difference - I was focused on the task. I may have done better in the second one simply because I had already practised."* (P-Colour-12)<br>*"I am not sure, because they were, in the end, quite similar."* (P-Room-03) | Lack of clarity in the preference environment and paying less attention to the surroundings. Moreover, focusing on the task makes no difference to the performance |

## Discussion

The results of this study indicate that learning environments aligned with individual spatial preferences are linked to improved performance on standardised cognitive tasks. Across all three environmental variables, most participants performed better in their preferred condition, with room proportion producing the strongest alignment rate (85.7%), followed by lighting (77.8%) and colour (75.0%). Overall, 79.2% of participants showed positive preference–performance alignment, substantially exceeding the 50% rate expected by chance. These findings suggest that the relationship between spatial preference and cognitive performance is systematic rather than random.

The qualitative themes provide explanatory depth for these quantitative patterns. Most of the "Aligned" group (79.2%) found that matching preferences and environment reduced distraction and anxiety, improving focus on the memory task. Participants described preferred settings as more comfortable, calming, or focused (like warm lights or cool colours), which aligns with lower cognitive load and better attention to the task. This aligns with P–E fit theory, which both predict that environments matching individual needs and expectations support better performance.

In contrast, the "Not aligned" group (20.8%) was characterised by the theme of Environmental Adaptation or Resistance. Participants often felt more comfortable in one condition but performed better in the non-preferred setting. Qualitative reports indicated that non-preferred environments sometimes led to increased effort, vigilance, or compensatory strategies. Preferred spaces were sometimes visually engaging or attractive but also more distracting, pulling attention from the memory task. Cognitive Affordances theory may explain why some participants reversed the expected preference–performance pattern, despite overall positive group results. For some participants, visually appealing or comfortable environments may have afforded relaxation, exploration, or mind-wandering, whereas less preferred environments may have encouraged greater vigilance, effort, or task-focused behaviour (Young & Cleveland, 2022). As a result, the environment that some participants preferred was not always the one that supported optimal memory performance.

Finally, the "Equal" group showed no clear performance differences across conditions, aligning with the theme of Environmental Neutrality. They reported weak or mixed preferences, focusing mainly on the task and paying little attention to environmental details. The learning task overshadowed environmental differences, resulting in similar performance across conditions. This suggests that for some learners, particularly those with strong task orientation or low sensitivity to spatial cues, the learning environment may serve as a relatively neutral backdrop, exerting minimal influence on performance.

Taken together, the quantitative alignment rates and the three qualitative themes support the view that the cognitive affordances of learning spaces are not uniform but are mediated by individual differences in preference, perceptual sensitivity, and attentional style (Gibson, 2014; Young & Cleveland, 2022). The same physical configuration can be a cognitively facilitative environment for one learner, whilst functioning as a source of distraction or indifference for another. This aligns with the proposition of Attention Restoration Theory that environments differentially support directed attention depending on individual states and environmental fit (Kaplan, 1995). These findings reinforce person–environment fit accounts of learning (Kristof-Brown et al., 2005), which similarly emphasise congruence between individual characteristics and environmental conditions as a driver of performance and wellbeing outcomes. These

findings underscore the value of adaptable, user-centred spatial designs that accommodate diverse cognitive needs, rather than relying on one-size-fits-all solutions.

Several limitations of this study should be acknowledged. This investigation relies on immersive virtual reality; although numerous studies show reliable results, there may be additional factors within the physical environment that are absent in the virtual setting. The cognitive memory tasks used are standardised; they may not fully represent the various aspects of academic performance seen in real-world learning environments. Future research should explore preference-performance alignment in physical educational settings. Using various designs and including a broader array of performance metrics, such as assessment grades and self-reported engagement, will enhance understanding.

## Conclusion

This study provides empirical evidence that learning spaces aligned with individual spatial preferences—specifically lighting, colour, and room proportions—are associated with improved cognitive task performance among university students. 79% of participants showed clear preference-performance alignment, with room proportions producing the largest effect (85%). These findings encourage reconsideration of the prevalent one-size-fits-all approach in educational settings. Moreover, the findings support a shift towards designing spaces that are more adaptable and genuinely aligned with individual cognitive needs and preferences.

The practical implications are considerable. Educational institutions should prioritise integrating individually adjustable lighting systems and providing varied zones, room sizes, and colours to accommodate a range of individual preferences. More broadly, designers should consider developing frameworks for learning spatial design that account for individual differences and cognitive needs, and that empower learners to meaningfully control learning spaces.

Future research should examine in depth the factors that may affect the spatial preference-performance relationship. This study offers a key foundation for the larger research effort, showing that the question of where learning occurs most effectively is not just about comfort but also about cognitive performance.

## Acknowledgements

This is part of a PhD at Lancaster University, funded by Taif University.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Grammarly was used to proofread and refine the manuscript's language. It was used to correct grammatical and spelling errors and rephrase statements for accuracy and clarity. The author further declares that, apart from Grammarly, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are original and derived from the appropriate and systematic conduct of the research.

## References

Aagaard, J. (2018). Magnetic and multistable: Reinterpreting the affordances of educational technology. *International Journal of Educational Technology in Higher Education*, *15*(1), 4.

Aagaard, J. (2018a). Magnetic and multistable: Reinterpreting the affordances of educational technology. *International Journal of Educational Technology in Higher Education*, *15*(1), 10.

Bandura, A. (1986). Social foundations of thought and action. *Englewood Cliffs, NJ*, *1986*(23–28), 2.

Barrett, P., Davies, F., Zhang, Y., & Barrett, L. (2015). The impact of classroom design on pupils' learning: Final results of a holistic, multi-level analysis. *Building and Environment*, *89*, 118–133. https://doi.org/https://doi.org/10.1016/j.buildenv.2015.02.013

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77–101.

Byers, T., Imms, W., & Hartnell-Young, E. (2018). Comparative analysis of the impact of traditional versus innovative learning environment on student attitudes and learning outcomes. *Studies in Educational Evaluation*, *58*, 167–177. https://doi.org/https://doi.org/10.1016/j.stueduc.2018.07.003

Choi, S., Guerin, D. A., Kim, H.-Y., Brigham, J. K., & Bauer, T. (2014). Indoor Environmental Quality of Classrooms and Student Outcomes: A Path Analysis Approach. *Journal of Learning Spaces*, *2*(2), 2013–2014.

Costa, A. R. (2025). Homing into school to create the ideal classroom: young people want to combine the best of home and school learning environments. *CoDesign*, *21*(1), 74–94. https://doi.org/10.1080/15710882.2024.2363925

Cronbach, L. J., & Snow, R. E. (1981). *Aptitudes and instructional methods: A handbook for research on interactions*. Ardent Media.

De Cooman, R., & Vleugels, W. (2022). Person–environment fit: theoretical perspectives, conceptualizations, and outcomes.

Fraser, B. J., & Fisher, D. L. (1983). Student achievement as a function of person-environment fit: A regression surface analysis. *British Journal of Educational Psychology*, *53*(1), 89–99. https://doi.org/10.1111/j.2044-8279.1983.tb02538.x

Gibson, J. J. (2014). *The ecological approach to visual perception: classic edition*. Psychology press.

Greeno, J. G. (1994). Gibson's affordances.

Jelić, A., Tieri, G., De Matteis, F., Babiloni, F., & Vecchiato, G. (2016). The enactive approach to architectural experience: A neurophysiological perspective on embodiment, motivation, and affordances. *Frontiers in Psychology*, *7*, 481.

Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology*, *15*(3), 169-182. https://doi.org/https://doi.org/10.1016/0272-4944(95)90001-2

Kariippanon, K. E., Cliff, D. P., Lancaster, S. L., Okely, A. D., & Parrish, A.-M. (2018). Perceived interplay between flexible learning spaces and teaching, learning and student wellbeing. *Learning Environments Research*, *21*(3), 301–320. https://doi.org/10.1007/s10984-017-9254-9

Kiverstein, J., & Artese, G. F. (2024). The experience of affordances in an intersubjective world. *Topoi*, *43*(1), 187–200.

Kozhevnikov, M., Evans, C., & Kosslyn, S. M. (2014). Cognitive Style as Environmentally Sensitive Individual Differences in Cognition: A Modern Synthesis and Applications in Education, Business, and Management. *Psychological science in the public interest*, *15*(1), 3–33. https://doi.org/10.1177/1529100614525555

Kristof-Brown, A. L., Zimmerman, R. D., & Johnson, E. C. (2005). Consequences of individuals'fit at work: A meta-analysis of person–job, person–organization, person–group, and person–supervisor fit. *Personnel Psychology*, *58*(2), 281–342.

Rosan Bosch Studio. (n.d.). *Vittra School Telefonplan. Retrieved*. Retrieved 8/2026 from https://www.rosanbosch.com/project/vittra-school-telefonplan

Rose, D. H., & Strangman, N. (2007). Universal Design for Learning: meeting the challenge of individual learning differences through a neurocognitive perspective. *Universal Access in the Information Society*, *5*(4), 381–391. https://doi.org/10.1007/s10209-006-0062-8

Russell, J. A., & Ward, L. M. (1982). Environmental psychology. *Annual review of psychology*, *33*(1), 651–689.

Young, F., & Cleveland, B. (2022). Affordances, Architecture and the Action Possibilities of Learning Environments: A Critical Review of the Literature and Future Directions. *Buildings*, *12*(1), 76. https://www.mdpi.com/2075-5309/12/1/76

Zhang, Z., Zhang, W., Zhang, S., Chen, Y., Wang, X., Fujii, Y., & Furuya, N. (2025). Person-environment fit theory in built environment: a scoping review. *Journal of Asian Architecture and Building Engineering*, *24*(4), 2447–2463. https://doi.org/10.1080/13467581.2024.2373824

**Contact emails:** w-aljehani@hotmai.com
w.aljehani@lancaster.ac.uk

# Designing and Facilitating External Partner-Based Pedagogy in Higher Education: Implications for Authentic Learning, Student Engagement, and Perceived Relevance

Rachel Masson, University of Glasgow, United Kingdom
Ana Da Silva Costa, University of Glasgow, United Kingdom
Jed Lennox, University of Glasgow, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Higher Education institutions are under increasing pressure to design learning experiences that engage students and demonstrate relevance to professional and societal contexts. Traditional classroom-based pedagogies have been criticised for limiting authentic learning and reinforcing a divide between academic knowledge and professional practice. External partner–based pedagogy, in which students collaborate with industry, community, or professional partners, offers a means of bridging this gap. Grounded in experiential and constructivist learning theories, such approaches require educators to act as designers, facilitators, and boundary-spanners between institutional and external contexts. This reflective study draws on the design and delivery of external partner–based learning activities across ten postgraduate life sciences courses in Scotland. Using reflection from staff observations, we examined student engagement during the course, the provision of authentic learning scenarios, and perceived relevance to learning outcomes. External partners contributed specialist expertise, real-world problem-solving activities, and employability-focused sessions, supporting a Trilateral Model of Benefits in which students, educators, and partners benefit reciprocally. Findings indicate that student motivation and attendance were highest when external sessions aligned clearly with module learning outcomes and assessments. When external speakers challenged students with real-life problems, engagement, participation, and depth of discussion increased. Exposure to diverse professional perspectives enhanced confidence, informed career decisions, and strengthened perceptions of curricular relevance. Overall, engaging external partners provides powerful opportunities for authentic learning and sustained engagement in postgraduate life sciences education. However, engagement becomes harder as cohort size increases, requiring integration, communication, and innovative approaches to scaling across programmes.

*Keywords:* external partners, higher education, authentic learning, employability

iafor
The International Academic Forum
www.iafor.org

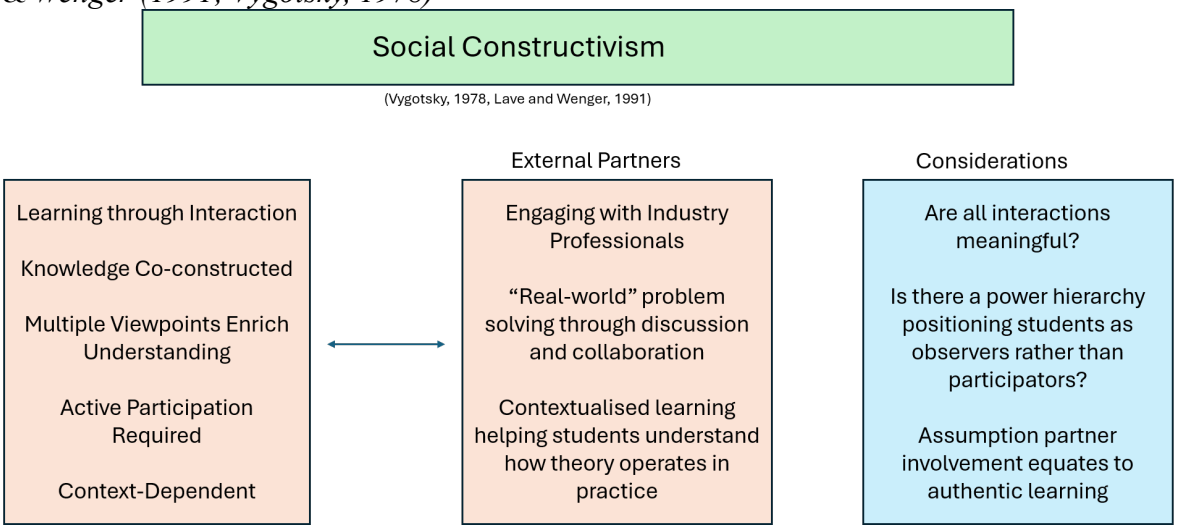
## Introduction

Higher Education (HE) institutions are increasingly challenged to design learning experiences that meaningfully engage students and demonstrate clear relevance to professional and societal contexts. Traditional classroom-based pedagogies have been criticised for their limiting authentic learning and reinforcing a separation between academic knowledge and professional practice. External partner–based pedagogy, where students collaborate with industry, community, or professional partners, has emerged as a promising approach to bridging this theory–practice divide (Marshall, 2018). Grounded in experiential and constructivist learning theories, such approaches require educators to adopt expanded roles as designers, facilitators, and boundary-spanners between institutional and external contexts (**Figure 1**) (Kolb, 2015; Piaget, 1973). Drawing on the design and facilitation of external partner–based learning activities implemented across ten postgraduate life sciences courses, this reflective study explores the implications for authentic learning, student engagement, and perceptions of curricular relevance. Through critical reflection on teaching practice across multiple disciplinary contexts within the life sciences, the paper contributes insights into the opportunities and challenges of scaling and sustaining partner-based pedagogy in contemporary HE.

The aims of this study are to provide a reflection on the benefits and shortcomings of engaging with external partners in the provision of post-graduate teaching in Higher Education for authentic learning, student engagement, and perceived relevance. We also aim to provide suggestions for those wishing to engage with external partners in different capacities with regards to level of engagement, good academic practice and recruitment.

**Figure 1**

*Social Constructivism Theory and Its Applications for This Study. Based on Studies by Lave & Wenger (1991; Vygotsky, 1978)*




## Methodology

British universities often engage with external speakers to provide specialist teaching and increase employability. Postgraduate (PGT) degrees are encouraged to provide real life-based scenarios to improve student engagement and employability options.

We reflect (Kane et al., 2004) on observations made by academic staff across ten different life sciences PGT courses in Scotland, and as former students. Using published literature on education, we inform and suggest applications to teaching in Higher Education PGT degrees.

We observed and compared student engagement during sessions and post-class, authentic learning scenario provision, and perceived relevance to student learning based on learning outcomes.

## Reflection

### How the Experience as a Former Student Influences and Affects Course Design

One author's previous experience in a collaborative programme as a student reported a consistently positive experience when engaging with external partners. This is based on external speakers bringing expert knowledge, non-traditional ways of presenting, and site-visits. This led to them feeling more comfortable in the knowledge they were receiving the best education possible, it helped inspire interests that would otherwise have remained undiscovered, and ultimately led them to where they are today with aspects of their education still actively influencing their own teaching practice and playing a role in their development as an educator. As such, we have looked into our experiences as educators in developing collaborative programmes.

### Selecting External Partners

The process of recruitment and selection of external partners is crucial from the first engagement. We must state that some external partners were contacted by university staff due to expertise, but some initiated contact with the academic staff prior to recruitment. As such, reflections and suggestions for engagement will be done for both circumstances.

#### *Recruitment of External Partners: Academic Staff Initiates Engagement*

Engagement with external partners is more commonly initiated by the academic staff involved in content development and teaching, in our experience. As a course is being designed, staff develop expectations of what is being taught to fulfil the intended learning outcomes (ILOs). As such, staff may want to engage with support from experts in the field who can provide a level of depth and understanding based on real-life scenarios.

Reasons for initial approach may include previous knowledge of their work, previous collaborations, networking, web searches for field of expertise. Our selection criteria was based on current role and current employment and relevant experience in the field of interest. We also welcome previous experience in delivering academic sessions.

#### *Recruitment of External Partners: External Partners Initiate Engagement*

Less frequently, external partners interested in collaborative work may contact us. Previously they have shown an interest in engaging with academia for networking and promoting roles within their industry. Some have queried about engaging at a deeper level due to skills gaps they found in recent graduates employed within their industry, and recommend content they deem useful for future employability. Should interests align and the external partner be properly vetted, we then initiate the next stage of selection.

### *Initial Meeting*

From our experience, we have found that it is important to have an initial meeting at least three months prior to teaching taking place. In this initial session, both parties (academic staff and external partners) agree in the content to be delivered in the context of the course and programme overall. It is important to ascertain that academic staff has control of the decision due to standards in academic practice needing to be maintained (see "*Ensuring teaching parity and academic standards"* below). Academic staff should highlight the ILOs for the session and the overall course to ensure all are met adequately. From previous experience and student reports, we have acknowledged the importance of presenting the intended assessments for the course to the external partner in medium and high engagement agreements (**Figure 2**). Where assessments were not previously discussed, students reported a disconnect from the sessions and, at times, a lack of understanding of where the session fitted into the overall course or its purpose as it did not contribute to the final assessment in any form.

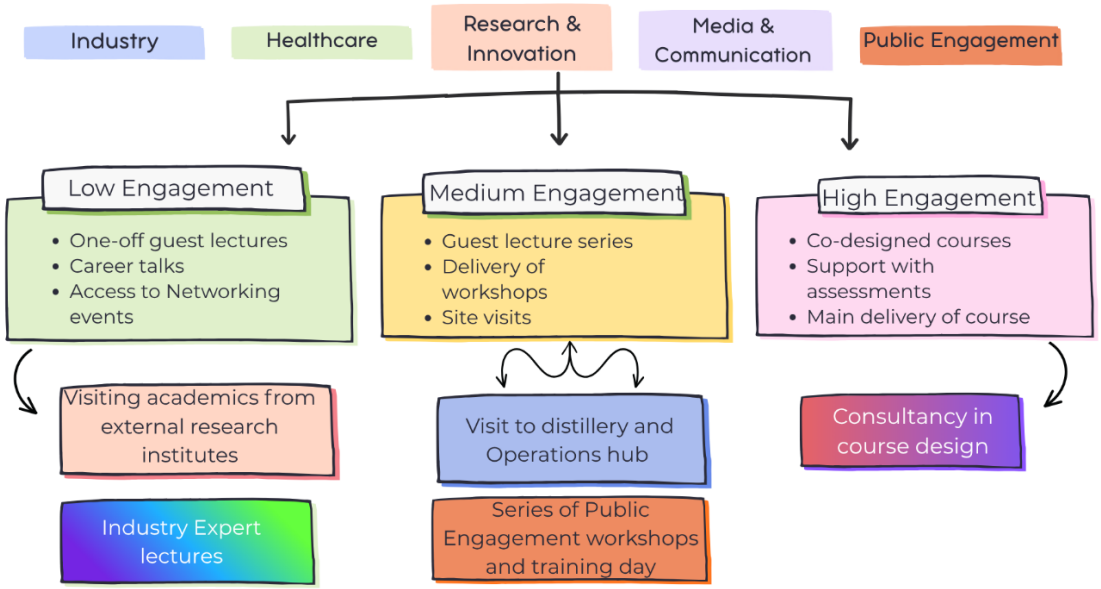
### *Levels of Engagement*

Our team established three levels of engagement from external partners through different collaborations (**Figure 2**). The most common is low engagement (**Figure 2**, green) where guests may provide a one-off lecture, a career talk, or access to a networking event. This is where we see visiting academics visiting from external research institutions or universities, or industry experts. In the courses provided by the authors, this was mostly done in industry-related, healthcare, and research and innovation courses.

Medium engagement (**Figure 2**, yellow) has external speakers provide 2 or more lectures, deliver workshops or provide site visits. As such, students have a higher level of engagement with the external partner and more opportunities for networking while querying about careers in the field. Our observations show that students have higher motivation in these sessions as they become more familiar with the speakers and have higher levels of confidence in their interactions. Examples of medium engagement in our courses were a series of Public Engagement workshops with an external partner, and a site visit to a distillery and operations hub following a lecture by an external speaker. This level of engagement was seen in courses related to industry field and Public Engagement.

The highest level of engagement (**Figure 2**, pink) we have experienced was where external partners supported the course design where these are mostly skills-based. Their advice helped create courses that promote employable skills, and create assessments that are similar and applicable to those experienced in professional careers, such as the creation of social media campaigns and interpretations of analytics from these. In some cases, the external partner may be the provider of the main content of the course in the presence of academic staff. The fields where we observed this were public engagement, and media and communication (**Figure 2**).

**Figure 2**
*Level of Engagement From External Partners and Examples of Contributions to Teaching*




## Ensuring Teaching Parity and Academic Standards

The academic team from relevant courses delivers most of the content on the course, but to ensure teaching parity and academic standards, also attends sessions provided by external partners. We engaged with partners who could provide personal and professional experience, deliver teaching and student-centred, active learning experience (Marshall, 2018). This also contributes to a Trilateral Model of Benefits (Zou et al., 2019), promoting reciprocal benefits where specialised teaching is provided, students learn from real-life scenarios, and external speakers become potential employers. Two courses added specialised sessions for students to work on CVs and interview skills with employers.

"When working in partnership, the awarding organisation retains responsibility for the academic standards of its awards, ensuring that the threshold standards for its qualifications are consistent with the relevant national qualification frameworks" (UKSCQA, 2023). As the qualifying organisation, it is the academic staff's responsibility to ensure there is compliance with the regulations.

As the qualifying organisation, we take the following steps to ensure academic standards and teaching parity: Follow the Quality Code for Higher Education; development of ILOs in line with Bloom's taxonomy guidelines (Adams, 2015) for PGT/SCQF degrees; set ILOs for individual sessions or multiple session contributions from External partners; conduct assessments (both formative and summative); attend all guest sessions to help field questions from students.

## Real-Life Scenario Enactment in Class

### *Example 1: Editorial Meetings*

In 2 of our courses in Media and Communications (**Figure 2**) we run weekly "Editorial Meetings." These aim to mimic a real-life scenario presented in journalism and social media

communications fields. In the academic sense, these help support the students with their assessments and provide formative feedback (Beaumont et al., 2011). In both courses, students are presented with the assessment on the first session of the course to allow time to become familiar with the assessment brief, and to start thinking about their topic of interest. In the first editorial meetings, students pitch their idea as they would a story or a campaign to a board of peers and both academic staff and the external partner. Everyone is invited to provide constructive feedback and suggest changes where needed. We emphasise the importance of everyone being able to participate while maintaining cordiality and professionalism. Staff monitors and intervenes should there be any potential conflict.

In subsequent meetings, the time is used to show what the student has been working on, and receive constructive feedback. The benefits of this format, as observed by staff, are that the real-life simulation presents a low-pressure environment and more students engage and participate than when feedback is provided mainly by a member of staff. Students receive support with assessments from staff, but also from peers. This also provides an opportunity to network with a potential employer while displaying learnt skills throughout the course.

### *Example 2: Public Engagement Activity*

In another course (in Public Engagement – **Figure 2**) the students work in collaboration with both an external partner and an internal partner over 10 weeks to design and deliver an activity for a public science event. Over this period, students have regular meetings with the festival organisers to pitch group activity ideas, and further to consider health and safety, audiences, materials needed, and so on. Members of academic staff are present in all meetings and communications to ensure activities stay aligned with the course aims and ILOs. It is worth noting that the assessments are the developed activity plan, and a reflection on the activity delivery to the general public while analysing feedback from participants.

The benefits we observed were that students develop valuable skills in event planning, science communication and project management, gain confidence in presenting complex scientific concepts, build professional networks, and reflect on the event and its effectiveness to support students in recognising and articulating skills they have developed.

## Benefits of the Different Types of Collaboration

Reflective analysis showed that allowing students to apply theoretical knowledge in practical settings deepened understanding and engagement where external speakers play a nuanced role in motivation and influencing educational outcomes. Our experience has shown that, when external speakers challenge students with real life-based problems, the cohort engaged greatly and was participative, actively discussing possible solutions. The class was more active and proactive in conversations while being exposed to various teaching styles, points of view, and opinions from experts in topics that they would not have otherwise experienced. This helped inform and influence decisions regarding their future, allowing them to communicate with people with varied and interesting academic journeys and backgrounds.

Studies have shown that guest speaker sessions can impact understanding of the real world and improve motivation (Pepple et al., 2025). Our observations show student motivation and attendance fluctuates under unknown circumstances, even when sessions are deemed compulsory when not connected to the assessments, therefore it is important for speakers to be

aware of the module requirements. Career talks are also seen as valuable networking opportunities.

### Future Directions

Our observations have greatly influenced our approach to teaching and engaging with external partners. The experience has been largely positive, and we will continue to do so. The next steps in our own learning are to investigate the student experience through surveys and determine whether our observations match their experience. We would also like to look into alumni employability to determine whether the presence and networking with external partners influenced the careers after graduation.

### Suggestions for Other Educators in Higher Education

Our experiences have helped shape our programmes and courses, from development to delivery. We review the collaborations yearly and take into account student feedback from student-staff meetings with the class representatives and from end-of-course surveys.

We found that support from external partners is essential in providing real-life experiences. We also saw more problem-solving exercises being presented to students. Students get a good picture of what real-life meetings look like in industry. We observed more engagement in class, with students posing more questions and networking for potential collaborations and employment. Attendance was higher on days where external partners were delivering sessions – but only if advertised to students in advance.

Collaborations support teaching in a positive manner, but agreements have to be made in advance. It is important to discuss what externals will deliver and how it connects to assessments and ILOs, and the overall course. There has to be a willingness on both parts to meet frequently and discuss problems as they come. Educators should be aware of GDPR and data management regulations (HESA, n.d.), especially in medium to high levels of engagement as student information should remain protected at all costs. We strongly recommend a member of staff is present in all sessions to monitor the content, ensure it connects to the assessments, learn so that assessments can be marked to a high standard of knowledge.

It is also advisable that staff make it clear to students that assessments queries should be sent to academic team as they have the oversight needed. Where it was not properly communicated in what capacity students should contact external partners or academic staff, we saw a mismatch of information being provided in places. As such, it is important that students know to contact external partners about content and employment, but assessments, extensions, illness and so on should only be communicated to academic staff.

## Conclusion

Engaging external partners in HE offers powerful opportunities for authentic learning and sustained student engagement. Our reflections across postgraduate life sciences courses show that when sessions align closely with learning outcomes, students connect theory with practice more deeply and develop a stronger sense of professional relevance. Collaboration with industry and community experts also broadens perspectives and nurtures confidence in career pathways. The difficulty of ensuring meaningful engagement scales with the size of the programme, smaller cohorts naturally able to have more contact time with external contributors

than larger cohorts. Therefore, to maximise impact, such activities should remain purposeful, clearly communicated, and inclusive. Future practice should explore how these approaches can be scaled sustainably, ensuring that the benefits of authentic, partner-based learning extend across programmes and disciplines.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that no AI or AI-assisted technologies have been used to generate, refine, or correct the content in the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Beaumont, C., O'Doherty, M., & Shannon, L. (2011). Reconceptualising assessment feedback: A key to improving student learning? *Studies in Higher Education*, *36*, 671–687. https://doi.org/10.1080/03075071003731135

Higher Education Statistics Agency. (n.d.). *Data protection*. Retrieved July 31, 2026, from https://www.hesa.ac.uk/about/regulation/data-protection

Kane, R., Sandretto, S., & Heath, C. (2004). An investigation into excellent tertiary teaching: Emphasising reflective practice. *Higher Education*, *47*, 283–310. https://doi.org/10.1023/B:HIGH.0000016442.55338.24

Kolb, D. A. (2015). *Experiential learning: Experience as the source of learning and development* (2nd ed.). Pearson Education.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Marshall, S. J. (2018). Internal and external stakeholders in higher education. In *Shaping the university of the future: Using technology to catalyse change in university learning and teaching* (pp. 77–102). Springer Singapore. https://doi.org/10.1007/978-981-10-7620-6_4

Pepple, D. G., Akaighe, G. O., Sambo, A., George-Aremu, O., Bosah, G., & Trollman, H. (2025). Using guest lectures to enhance student employability: Pedagogical considerations. *Cogent Education, 12*(1), 2452076. https://doi.org/10.1080/2331186X.2025.2452076

Piaget, J. (1973). *To understand is to invent: The future of education*. Grossman Publishers. https://unesdoc.unesco.org/ark:/48223/pf0000006133

The UK Standing Committee for Quality Assessment. (2023). *UK Quality Code for Higher Education 2018*. https://www.qaa.ac.uk/docs/qaa/quality-code/uk-quality-code-for-higher-education-2018.pdf?sfvrsn=4c19f781_31

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.

Zou, P., Sun, W., Hallowell, S. G., Luo, Y., Lee, C., & Ge, L. (2019). Use of guest speakers in nursing education: an integrative review of multidisciplinary literature. *Advances in Medical Education and Practice*, *10*, 175–189. https://doi.org/10.2147/AMEP.S196456

**Contact email:** ana.costa@glasgow.ac.uk

# The Early Childhood Education Crisis Across Socioeconomic Contexts: A Comparative Qualitative Study From Israel

Amit Rottman, The Max Stern Yezreel Valley College, Israel
Yoad Eliaz, The Max Stern Yezreel Valley College, Israel
Lior Gendelman, The Max Stern Yezreel Valley College, Israel

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Early Childhood Education (ECE) is critical for children's cognitive, emotional, and social development. In Israel, however, several indicators of ECE quality continue to rank among the lowest across OECD countries. Daycare centers operate under structural crisis conditions, marked by resource constraints, staff shortages, high workforce turnover, unstable regulation, low wages, and insufficient professional training. While previous research has documented these conditions primarily through quantitative indicators, this study offers a qualitative account of how the ECE crisis is experienced in the everyday work of daycare staff. Given the central role of high-quality ECE in narrowing socioeconomic gaps, the study compared daycare centers in two socioeconomically distinct localities: a disadvantaged, peripheral city and a middle-class kibbutz. Data collection combined semi-structured interviews, focus groups, participant observation, document analysis, and follow-up meetings across both settings. Findings reveal that the structural crisis characterizing Israeli ECE manifests in both settings, but in markedly different forms. In the disadvantaged urban setting, the crisis was experienced in an intensified, acute manner, compounding pre-existing social vulnerability. In the more affluent kibbutz, the same underlying complexities emerged under more privileged, "first-class" conditions, mitigated by strong local leadership, greater organizational flexibility, and closer coordination among professional supports. Overall, the comparison suggests that socioeconomic advantage does not eliminate the ECE crisis but reshapes its intensity and lived experience. This conclusion underscores the need to address national-level funding, supervision, and guidance, alongside local organizational and managerial practices.

*Keywords:* early childhood education, daycare crisis, socioeconomic inequality, qualitative research, Israel

iafor
The International Academic Forum
www.iafor.org

## Introduction

Early childhood education (ECE) is widely recognized as an important influence on children's development (Britto et al., 2017; Conceição et al., 2025; Daelmans et al., 2017) and as a means of reducing gaps created by social and economic inequality (Heckman, 2006, 2012; Solang et al., 2024). High-quality ECE settings are associated with cognitive, linguistic, social, and emotional gains that persist well beyond the preschool years. These gains are particularly important in low-income communities, where access to high-quality ECE could serve as a key mechanism for promoting equity (Solang et al., 2024).

In Israel, ECE for children from birth to age three operates largely outside the formal education system, through a fragmented structure of municipal, private, nonprofit, and cooperative daycare settings. This fragmented system is supervised and regulated by multiple government authorities and is characterized by persistently low-quality indicators relative to OECD countries (OECD, 2017; Vaknin, 2020). Although participation in early childhood programs in Israel is high, the system's low-quality structure limits ECE settings' ability to identify developmental challenges early and promote equity from the beginning of life. Studies examining the quality of Israeli ECE consistently report low standards: large group sizes, high child-to-caregiver ratios, chronic staff shortages, high employee turnover, insufficient professional training, and limited attention to children's emotional needs (Koren-Karie et al., 2005; Vaknin, 2020; Zano, 2020). These conditions point to a systemic crisis in the provision of quality ECE. While this crisis characterizes the Israeli ECE system as a whole, its everyday manifestations and long-term meanings may differ across distinct socioeconomic contexts.

Against this background, the present study compared two qualitative case studies conducted in contrasting socioeconomic settings. The first case examined the everyday work of caregivers and managers in a socioeconomically disadvantaged, peripheral city. The second was conducted in a middle-class kibbutz community. The comparison was guided by two questions: (1) How is the ECE crisis experienced in the everyday work of daycare staff and managers across socioeconomically advantaged and disadvantaged settings? and (2) What structural and symbolic conditions shape the form and severity of the crisis across these settings? By placing the two cases side by side, the study aimed to move beyond system-level quantitative indicators of the ECE crisis to develop a situated understanding of how socioeconomic inequality shapes the everyday realities of early childhood education and care.

## Literature Review

### The Significance of Early Childhood Education

Early childhood experiences have long-term effects (Parise et al., 2023; Short et al., 2019; Xu et al., 2025). Infants and toddlers, especially from birth to age three, build their cognitive, emotional, and social worlds (Britto et al., 2017; Conceição et al., 2025; Daelmans et al., 2017; Zontag et al., 2020). Developmental gaps that emerge in early years tend to accumulate over time (Kuijper & Johnstone, 2019). Research has therefore highlighted the importance of investments made during the first years of life (Harrison et al., 2012; Heckman, 2006, 2012; Moore & West, 2016).

Participation in high-quality ECE programs has been shown to significantly enhance young children's development, while also improving long-term educational, economic, and health outcomes (Denham & Brown, 2010; Zontag et al., 2020). However, considerable variation

exists between daycare centers, influencing their capacity to create high-quality care. This variation is shaped by structural factors, such as child-to-caregiver ratios and group size (Bowne et al., 2017), as well as by caregivers' training, education, and professional experience (Burchinal et al., 2016). The quality of daycare centers improves when caregivers have higher levels of formal education and professional training (Lin & Magnuson, 2018) and when they understand early socio-emotional development and respond sensitively to young children's emotional cues (Sabol & Pianta, 2012).

Taken together, research indicates that children's development and outcomes are strongly influenced by the quality of daycare environments and the nature of caregiver-child interactions (Muennig et al., 2011; Phillips et al., 2011; Solang et al., 2024; Von Suchodoletz et al., 2023). Investment in ECE yields multiple educational, economic, emotional, and social benefits, some of which are expressed later in life. These benefits are particularly critical for children at risk. Research shows that ECE has a significantly more positive effect on toddlers from low socioeconomic backgrounds than on their more affluent peers. Thus, when made universally accessible, quality ECE has the potential to reduce developmental disparities and promote social equality (Solang et al., 2024).

## Early Childhood Education in Israel

In Israel, education for children under the age of three is provided through several types of institutions, including daycare centers, family daycares, toddler centers, and private kindergartens. Daycare centers, which are the focus of the present paper, are supervised frameworks operated by national organizations, nonprofit associations, local authorities, or private providers. They are commonly classified into three categories: small settings (7–20 children), medium-sized settings (21–50 children), and large settings (51 children or more) (Madhala et al., 2021).

Until 2022, formal responsibility for daycare centers rested primarily with the Ministry of Labor, Welfare, and Social Services. In early 2022, responsibility for pedagogical supervision, staff training, and the development of daycare facilities began to be transferred to the Ministry of Education. However, other areas remain under the authority of different government bodies: the Ministry of Welfare and Social Services retains responsibility for toddlers with disabilities and children defined as being at risk, while the Ministry of Economy administers subsidies and eligibility for publicly supported placements (Rabinowitz, 2023). National daycare organizations and local authorities also continue to play significant roles in the management and operation of settings (Zano, 2020). Consequently, despite the transfer of formal educational responsibility to the Ministry of Education, the system remains divided among multiple governmental and organizational actors.

This fragmented structure is closely related to multiple problems of regulation, funding, and professionalization, which are common in the Israeli ECE system. Although participation in early childhood settings in Israel is high, the quality of provision has repeatedly been described as low (Vaknin, 2020). Public investment in ECE remains among the lowest in OECD countries (OECD, 2017). In addition, regulation has historically been partial and uneven, particularly in privately operated settings, where uniform requirements concerning staff qualifications, training, and suitability have not always been enforced (Madhala et al., 2021; Rabinowitz, 2019). Even in supervised daycare centers, caregivers may begin working without prior professional training, formal certifications, or a systematic assessment of their personal suitability for working with infants and toddlers. Moreover, the basic training provided by

supervising organizations is often insufficiently responsive to the distinctive developmental and emotional needs of young children (Zano, 2020).

Studies of Israeli daycare centers have identified multiple structural and pedagogical difficulties, including large group sizes, high child-to-caregiver ratios, inadequate professional training, limited attention to children's individual emotional needs, and shortages of qualified and experienced staff (Koren-Karie et al., 2005; Vaknin, 2020; Zano, 2020). These problems are intensified by low wages, poor employment conditions, difficulties in recruitment and retention, and weak or inconsistent supervision mechanisms (Rabinowitz, 2023). These conditions have contributed to an ongoing crisis in the ECE sector, expressed not only in measurable indicators such as staffing ratios and workforce turnover, but also in the limited capacity of daycare centers to provide stable, developmentally appropriate, and emotionally responsive care (Vaknin, 2020; Zano, 2020).

These manifestations are documented in the literature mainly from a structural perspective, with quantitative indicators such as staffing ratios, wage levels, and turnover rates (Rabinowitz, 2023; Vaknin, 2020; Zano, 2020). While these measures capture the wide scope of the Israeli ECE crisis, they offer limited insight into how this crisis is experienced in the everyday work of daycare staff and managers. The present study seeks to deepen the understanding of the ECE crisis by examining its concrete, lived manifestations in daycare centers. Informed by the claim that socioeconomic status and resources may shape the concrete quality and experience of children inside educational environments (Hatfield et al., 2014; Wei et al., 2021), it compares the consequences of the ECE crisis across two contrasting socioeconomic contexts.

## Methodology

The study employed a comparative qualitative case study design. Two daycare settings, located within the same geographic region in northern Israel, were selected to represent different socioeconomic backgrounds. Case 1 was conducted in a socioeconomically disadvantaged, peripheral city, and Case 2 in a middle-class kibbutz. Findings from both cases were then analyzed jointly through a cross-case comparative lens.

### Research Sites and Participants

#### *Case 1: A Disadvantaged Peripheral City*

The first case study was conducted across daycare centers operating in a small city with approximately 30,000 residents. According to the Israeli Central Bureau of Statistics (2021), the city is ranked in socioeconomic cluster 5 out of 10.

The city's ECE system comprises medium- and large-sized supervised daycare centers, managed by various private and nonprofit providers. A municipal ECE team oversees these providers and provides them with professional guidance and support. Members of the municipal ECE team explained that a significant number of children enrolled in the city's supervised daycare centers come from lower-middle and working-class families, many of whom face various socioeconomic challenges (such as single-parent households, immigrant status, or large family size).

Data collection focused on five daycare centers, all classified as medium- to large-sized. Each center was organized into two or three separate age groups. Each group had one lead caregiver

and two to three additional staff members. Each daycare center also had a general manager who was responsible for administrative and financial management, human resources, parent relations, and so on.

### *Case 2: A Middle-Class Kibbutz*

The second case study was conducted in a kibbutz community, also located in northern Israel, but unlike Case 1, classified as middle- to upper-middle-class in socioeconomic terms. The kibbutz's ECE system was characterized by a cooperative communal structure, and the caregiving workforce was employed directly by the kibbutz rather than by an external municipal body or nonprofit providers.

Children from birth to age three were cared for across three separate age-based settings: an infant center (birth to one year), a toddler group (one to two years), and a young kindergarten group (two to three years). Each setting had a lead caregiver (or, in the case of the young kindergarten group, a certified kindergarten teacher), along with three to four additional caregivers. The system as a whole was headed by a manager who served as a senior member of the kibbutz management and was responsible for all ECE-related matters.

The child population in the kibbutz's daycare settings comprised two distinct groups. The first consisted of children whose parents lived in the kibbutz and were part of its community. The second consisted of children whose parents resided in nearby localities, and they paid substantial fees to enroll their children in the kibbutz's ECE system. These parents were, in most cases, career-oriented individuals seeking a setting that would provide their children with a calm environment for extended daily hours, while also offering a solution for workdays that fell during the formal education system's holidays. Kibbutz-based ECE systems are widely perceived as addressing both of these needs. They are often regarded in Israel as high-quality frameworks, known for values-based education, frequent outdoor activities in nature, and considerable resources. In addition, they offer extended hours of operation during school holidays and vacations.

## Data Collection and Analysis

Chronologically, the study was conducted in two phases: first in the urban locality (2020–2021), and later in the kibbutz (2024–2025). This sequence reflects the study's broader rationale: having documented the ECE crisis in a disadvantaged, working-class setting, the research team sought to examine how the same structural crisis is experienced in a middle-class community to identify points of similarity and difference between the two settings.

Data collection combined participant observations, semi-structured in-depth interviews, focus groups, and document analysis, supplemented by follow-up contact via phone calls and Zoom meetings. Participants included ECE municipal administrators, the kibbutz head of the ECE system, daycare managers, and caregivers. Table 1 and Table 2 summarize the data sources for Case 1 and Case 2, respectively.

**Table 1**
*Data Collection Activities, Case 1 (Disadvantaged City)*

| Data Source | Interviews | Focus Groups | Additional Meetings |
|---|---|---|---|
| Daycare managers | 5 | - | - |
| Caregivers | - | 5 | - |
| Municipal ECE administrators | 2 | - | - |
| Follow-up contacts (phone/Zoom) | - | - | 6 |
| **Total** | **7** | **5** | **6** |

**Table 2**
*Data Collection Activities, Case 2 (Middle-Class Kibbutz)*

| Data Source | Interviews | Focus Groups | Additional Meetings |
|---|---|---|---|
| Municipal ECE administrator | 1 | - | 1 |
| Head of the kibbutz ECE system | 1 | - | 2 |
| Lead caregivers | - | 1 | - |
| Caregivers | - | 2 | - |
| **Total** | **2** | **3** | **3** |

In both cases, interviews and focus groups were recorded, transcribed, and analyzed thematically, first within each case and then across cases. The cross-case comparative stage focused on identifying themes that were shared across the two settings, as well as themes that diverged in form, intensity, or meaning.

### Ethical Considerations

Both case studies received approval from the institutional ethics committee prior to data collection. Participation was voluntary and based on informed consent. To protect participants' anonymity, all identifying details have been removed. In addition, participants are identified by role only (e.g., “head of the kibbutz ECE unit,” “lead caregiver”) rather than by name.

## Findings – A Shared Crisis, Different Conditions

Across both cases, participants described a shared set of structural difficulties - staff shortages, demanding workloads, inconsistent professional training, and tensions in relationships with parents. The shared problems confirmed that the crisis documented in the disadvantaged city was not unique to that setting. At the same time, the two cases diverged in the intensity, visibility, and lived meaning of these problems, as detailed below.

### Staff Shortage and Workforce Instability

In both settings, participants described a persistent difficulty recruiting and retaining qualified caregivers. In the disadvantaged city, this shortage was described as acute and persistent, forcing managers to operate with higher child-to-caregiver ratios for extended periods and to absorb the resulting workload personally. In the kibbutz, staff shortages were buffered by the kibbutz's ability to draw on a wider pool of potential substitute caregivers, including, for example, older kibbutz members and women with disabilities seeking supported employment.

The head of the kibbutz's ECE system mentioned, for instance, the support she receives from the kibbutz's seniors' club: *"They are an integral part of the work routine – of the children's daily life, of ceremonies, of everything. It gives you that extra something."*

### Wages and Employment Structure

Caregivers in both settings were employed at or near minimum wage, reflecting a national pattern in which ECE caregiving is compensated as low-skilled labor regardless of the socioeconomic context in which it takes place. In the disadvantaged city, low wages compounded broader financial precarity and were repeatedly cited by caregivers and managers as a central driver of turnover. For instance, one daycare manager noted the mismatch between the demanding nature of the job and the minimal compensation: *"There is a huge gap between what we do and the responsibility we carry, and the wage, that fails to compensate for it."*

A lead caregiver at another daycare also echoed this perceived gap: *"to give everything for the benefit of the children and receive nothing [for yourself]."*

In the kibbutz, employment was structured differently: caregivers were employed directly by the kibbutz rather than through an external municipal or private body, and the head of the kibbutz's ECE system held discretion to authorize additional compensation or benefits (for example, in the form of additional paid vacation days or holiday gifts) and to absorb budgetary deficits at the community level. This arrangement did not eliminate low base wages. Nevertheless, it reduced the sense of employment precarity and unfairness described by caregivers in the disadvantaged city.

### Trust and Tension Between Caregivers and Parents

In both cases, caregivers described their relationships with parents as a possible source of professional strain, shaped by parents' anxiety about entrusting very young children to the care of others and by caregivers' awareness of being closely, and at times suspiciously, observed. In the disadvantaged city, caregivers described a heightened sense of being viewed with distrust, including concerns about being perceived as potentially neglectful or abusive. This sense of suspicion added an emotional burden to already demanding daily routines.

In the kibbutz, caregivers likewise described the need to actively maintain parents' trust and, at times, to apologize or justify their decisions or actions. Nevertheless, this tension was described as occurring within a broader context of communal familiarity and long-term relationships between families and daycare staff, which tempered its intensity. In addition, tensions with parents were tempered by the significant role of the head of the kibbutz's ECE system, who maintained meaningful relationships with both staff members (a matter that will be further elaborated below) and parents. The head of the ECE system's mode of operation created a

bridge between caregivers and parents when needed and, no less important, provided ongoing support for caregivers in their interactions with parents.

### Professional Training and the Coordination of Guidance

Consistent with the broader difficulties of the Israeli ECE system described throughout the literature review, access to professional training and pedagogical guidance was inconsistent and difficult to obtain in both settings. In the disadvantaged city, the lack of sufficient training was accompanied by overlapping and at times contradictory guidance from multiple supervising bodies, which produced confusion about expectations and reduced the practical value of the support that was available.

In the kibbutz, on the other hand, caregivers had access to a more coordinated support structure, including a pedagogical advisor affiliated with the Kibbutz Movement (*Hatnua HaKibbutzit*), working alongside the head of the kibbutz's ECE system. While the kibbutz's ECE system was also characterized by insufficient training for the caregivers and lack of formal education, the degree of coordination among the professionals involved was markedly higher than in the disadvantaged city. In addition, the kibbutz had the resources to send caregivers for further professional training and education.

### Leadership and a Sense of Belonging

A theme unique to the kibbutz case was the centrality of a single, stable leadership figure – the head of the kibbutz's ECE system – who was described by caregivers as a source of continuity, fairness, and personal investment in their well-being. Several caregivers reported that their decision to work, or continue working, in the kibbutz's daycare settings was tied specifically to their relationship with this individual, describing a sense of belonging to a small, family-like organizational unit rather than to an anonymous municipal system.

This was echoed in a focus group with caregivers, one of whom recalled: *"If we don't feel well, she [the head of the ECE system] tells us immediately to go home, she'll find someone else [to replace us]."* In response, another caregiver explained how this responsiveness sustained her motivation: *"It gives you the motivation to keep working. It is the kind of support that keeps you from burning out."*

No equivalent figure or organizational structure was described in the disadvantaged city, where responsibility for daycare oversight was distributed across a municipal department without a comparably personal or stable point of leadership.

## Discussion and Conclusion

Taken together, the findings support the view that the structural crisis affecting Israeli ECE – staff shortage, low wages, inconsistent training, and strained caregiver-parent trust – is not confined to socioeconomically disadvantaged settings but extends into middle-class communities as well. This finding extends previous accounts of the Israeli ECE crisis (Madhala et al., 2021; Rabinowitz, 2019; Vaknin, 2020), which have documented the scope and severity of the crisis primarily through quantitative data and system-level indicators. Building on this body of research, the present study examined how the ECE crisis is enacted and experienced in the everyday practice of caregivers and managers across different socioeconomic environments.

While the ECE crisis in Israel appears to be widespread, the comparison shows that socioeconomic status and resources shape the way it is experienced and functions in practice. As might be expected from studies showing that the quality of ECE is associated with neighborhood resources, community socioeconomic conditions, and program funding (Hatfield et al., 2014; Wei et al., 2021), the kibbutz's greater organizational, financial, and social resources appeared to buffer the effects of the ECE crisis. A wider pool of potential substitute caregivers reduced the practical impact of staff shortage, and discretionary authority over compensation reduced the sense of employment precarity. In addition, better-coordinated professional support increased the practical value of available training, and a stable, personally invested leadership figure fostered a sense of belonging that was absent in the disadvantaged city. These mechanisms suggest that the form and severity of the ECE crisis are shaped by the organizational and communal resources available. Thus, the findings offer a complementary lens for understanding how social inequality operates within ECE systems.

In sum, the comparative findings indicate that socioeconomic advantage does not eliminate the ECE structural crisis but substantially reshapes its intensity and lived expression through organizational buffering mechanisms. For policy, this suggests that interventions addressing the ECE crisis need to attend not only to system-wide structural factors such as wages and staffing ratios, but also to the organizational and community-level resources – leadership continuity, staffing flexibility, and coordinated professional support – that appear capable of buffering their effects. For research, the study suggests that future work on the ECE workforce would benefit from deliberately comparative designs spanning multiple socioeconomic contexts, in order to move beyond aggregate, system-level indicators toward a fuller understanding of how inequality shapes the everyday experience of early childhood education.

### Limitations

Several limitations should be noted. First, the two cases differ in the volume and depth of data collected, with the kibbutz case comprising a smaller number of interviews and focus groups than the disadvantaged-city case. This asymmetry may limit the extent to which the full range of experiences in the kibbutz setting has been captured. Second, data for the two cases were collected during different periods (2020–2021 and 2024–2025), a window during which the governance of Israeli ECE underwent institutional changes. Future comparative research would benefit from collecting data within the same timeframe to more clearly isolate the role of socioeconomic context from broader temporal and regulatory shifts. Third, as with any qualitative case study, the findings are not intended to be statistically generalizable to all socioeconomically advantaged or disadvantaged ECE settings in Israel, but rather to offer analytically transferable insight into the mechanisms through which socioeconomic context shapes the ECE crisis.

## Acknowledgements

The first case study reported in this paper was supported by a research grant from the MOFET Institute.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, the authors used Claude (Anthropic) to assist with organizing and drafting sections of the text based on the authors' own prior writing. In addition,

Grammarly was used to assist with language editing. All scientific content, analysis, and interpretation remain the sole responsibility of the authors, who reviewed and edited all AI-assisted text and take full responsibility for the content of the publication.

## References

Bowne, J. B., Magnuson, K. A., Schindler, H. S., Duncan, G. J., & Yoshikawa, H. (2017). A meta-analysis of class sizes and ratios in early childhood education programs: Are thresholds of quality associated with greater impacts on cognitive, achievement, and socioemotional outcomes? *Educational Evaluation and Policy Analysis*, *39*(3), 407–428. https://doi.org/10.3102/0162373716689489

Britto, P. R., Lye, S. J., Proulx, K., Yousafzai, A. K., Matthews, S. G., Vaivada, T., & MacMillan, H. (2017). Nurturing care: Promoting early childhood development. *The Lancet, 389*(10064), 91–102. https://doi.org/10.1016/S0140-6736(16)31390-3

Burchinal, M., Zaslow, M., Tarullo, L., Votruba-Drzal, E., & Miller, P. (2016). Quality thresholds, features, and dosage in early care and education: Secondary data analyses of child outcomes. *Monographs of the Society for Research in Child Development*, *81*(2), 1–126.

Central Bureau of Statistics. (2021). *Characterization of geographical units and their classification according to the socio-economic level of the population 2019*. https://www.cbs.gov.il/he/mediarelease

Conceição, C., Cadima, J., Camacho, A., & Alves, D. (2025). Testing the effectiveness of a social and emotional skill program for preschool children. *Early Childhood Education Journal*, 1–15. https://doi.org/10.1007/s10643-025-01909-8

Daelmans, B., Darmstadt, G. L., Lombardi, J., Black, M. M., Britto, P. R., Lye, S., & Richter, L. M. (2017). Early childhood development: The foundation of sustainable development. *The Lancet, 389*(10064), 9–11. https://doi.org/10.1016/S0140-6736(16)31659-2

Denham, S. A., & Brown, C. (2010). "Plays nice with others": Social-emotional learning and academic success. *Early Education and Development, 21*(5), 652–680.

Harrison, L., Goldfeld, S., Metcalfe, E., & Moore, T. (2012). *Early learning programs that promote children's developmental and educational outcomes*. AIHW.

Hatfield, B. E., Lower, J. K., Cassidy, D. J., & Faldowski, R. A. (2014). Inequities in access to quality early care and education: Associations with funding and community context. *Early Childhood Research Quarterly*, *30*, 316–326. https://doi.org/10.1016/j.ecresq.2014.01.001

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900–1902. https://doi.org/10.1126/science.1128898

Heckman, J. J. (2012). Invest in early childhood development: Reduce deficits, strengthen the economy. *The Heckman Equation, 7*.

Koren-Karie, N., Sagi-Schwartz, A., & Egoz-Mizrachi, N. (2005). The emotional quality of childcare centers in Israel: The Haifa study of early childcare. *Infant Mental Health Journal, 26*(2), 110–126. https://doi.org/10.1002/imhj.20039

Kuijper, B., & Johnstone, R. A. (2019). The evolution of early-life effects on social behaviour—why should social adversity carry over to the future? *Philosophical Transactions of the Royal Society B*, *374*(1770), 20180111.

Lin, Y. C., & Magnuson, K. A. (2018). Classroom quality and children's academic skills in child care centers: Understanding the role of teacher qualifications. *Early Childhood Research Quarterly*, *42*, 215–227. https://doi.org/10.1016/j.ecresq.2017.10.003

Madhala, S., Gal, J., Shami, S., & Elizur, A. (2021). *Enrollment of Arab children in supervised daycare*. Taub Center (Hebrew). https://www.taubcenter.org.il/wp-content/uploads/2021/11/EnrollmentinSupervisedDaycareHeb-1.pdf

Moore, T., & West, S. (2016). Early intervention and the first thousand days of child development. *Developing Practice*, *44*, 16–24.

Muennig, P., Robertson, D., Johnson, G., Campbell, F., Pungello, E. P., & Neidell, M. (2011). The effect of an early education program on adult health: The Carolina Abecedarian Project randomized controlled trial. *American Journal of Public Health*, *101*(3), 512–516. https://doi.org/10.2105/AJPH.2010.200063

OECD. (2017). *Starting Strong 2017: Key OECD Indicators on Early Childhood Education and Care*. OECD Publishing. https://www.oecd.org/education/starting-strong-2017-9789264276116-en.htm

Parise, L. F., Burnett, C. J., & Russo, S. J. (2023). Early life stress and altered social behaviors: A perspective across species. *Neuroscience Research*, *211*, 65–74. https://doi.org/10.1016/j.neures.2023.11.005

Phillips, D. A., Fox, N. A., & Gunnar, M. R. (2011). Same place, different experiences: Bringing individual differences to research in child care. *Child Development Perspectives*, *5*(1), 44–49. https://doi.org/10.1111/j.1750-8606.2010.00155.x

Rabinowitz, M. (2019). *Daycare frameworks for early childhood and government supervision over them* (Hebrew). The Knesset, Research and Information Center.

Rabinowitz, M. (2023). *Early childhood reform in the education system* (Hebrew). The Knesset, Research and Information Center.

Sabol, T., & Pianta, R.C. (2012). Recent trends in research on teacher-child relationships, *Attachment and Human Development, 14*(3), 213–231.

Short, A. K., Bolton, J. L., & Baram, T. Z. (2019). Mechanisms by which early-life experiences promote enduring stress resilience or vulnerability. In A. Chen (Ed.), *Stress resilience* (pp. 165–180). Academic Press.

Solang, D. J., Yunus, M., Rukmini, B. S., Kurnia, K., & Umam, K. (2024). The impact of early childhood education on long-term outcomes. *The Journal of Academic Science*, *1*(4), 350–361. https://doi.org/10.59613/7pvaa271

Vaknin, D. (2020). *Early childhood education and care in Israel compared to the OECD*. Taub Center. https://www.taubcenter.org.il/wp-content/uploads/2021/02/earlychildhood2020en.pdf

Von Suchodoletz, A., Lee, D. S., Henry, J., Tamang, S., Premachandra, B., & Yoshikawa, H. (2023). Early childhood education and care quality and associations with child outcomes: A meta-analysis. *PLoS One*, *18*(5), e0285985.

Wei, W.S., McCoy, D.C., Busby, A.K., Hanno, E.C., & Sabol, T.J. (2021). Beyond neighborhood socioeconomic status: Exploring the role of neighborhood resources for preschool classroom quality and early childhood development. *American Journal of Community Psychology, 67*(3-4), 470–485. https://doi.org/10.1002/ajcp.12507

Xu, Y., Cage, J., Chen, C. C., Zhang, F., & Xia, Y. (2025). Assessing young children's social-emotional development through the lens of social-cultural ecological systems. *Early Childhood Education Journal*, 1–11. https://doi.org/10.1007/s10643-025-01857-3

Zano, G. (2020). Attitudes of caregivers in the infant class in the day-care center. *Early Child Development and Care*, *190*(9), 1396–1410. https://doi.org/10.1080/03004430.2018.1538134

Zontag, N., Navon, Y., Vaknin, D., Bowers, L., Blank, C., & Shavit, Y. (2020). *Early Childhood Education in Israel and Academic Achievement*. Taub Center. https://www.taubcenter.org.il/wp-content/uploads/2020/11/educationalframeworksforyoungchildrenandacademicachievementeng1.pdf

**Contact email:** amitr@yvc.ac.il

# Visualizing the Invisible: Co-creating an AI-Assisted Tactile Picture Book With a Congenitally Blind Student

Chu-Yu Cheng, Southern Taiwan University of Science and Technology, Taiwan
Kuan-Ju Ho, Southern Taiwan University of Science and Technology, Taiwan

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This study explores how congenitally blind students perceive and imagine biological forms through co-creating an AI-assisted tactile picture book. Our research laboratory specializes in developing educational materials for visually impaired students. After completing several projects helping students understand biological characteristics, we became curious: how do students without visual experience conceptualize living creatures? We used AI as a creative tool and aesthetic design as a communication channel to externalize a blind student's imaginative world. The visually impaired (VI) student, along with design professionals, collaborated using Google Labs Mixboard to co-create a picture book about urban insects. The 15-year-old congenitally blind student developed characters, invented storylines, and orally described scenes. Inspired by Chinese martial arts novelist Louis Cha Leung-yung, he created a ten-legged cockroach protagonist with martial arts-style actions and settings. Prompts derived from his descriptions were input into Mixboard for image generation, with designers refining the outputs. The final 24-page book was produced using swell paper printing technology, creating tactile graphics accompanied by Chinese Braille. This collaborative project reveals that VI students' biological understanding often stems from verbal descriptions and tactile toys, resulting in imaginative yet scientifically inaccurate conceptions. The methodology demonstrates how AI-assisted co-creation can bridge imagination and reality while providing valuable insights into alternative cognitive frameworks and celebrating the boundless creativity of students with visual impairments. Its contribution is an account of a working method and a set of transferable principles, not a measure of learning effectiveness.

*Keywords:* practice-based design research, congenital blindness, tactile picture book, generative image models, authorship

iafor
The International Academic Forum
www.iafor.org

## Introduction

Generative AI is now used in classrooms, including in special education (Holmes & Tuomi, 2022; Khlaif et al., 2026). A language model can turn what a blind author imagines into a picture, and digital fabrication can turn that picture into an object he can touch. This was not possible before (Tanaka & Ochiai, 2023). However, most special education still treats blind students as people who receive information that others prepare for them (Khlaif et al., 2026).

This study began during earlier work on tactile teaching materials. A 15-year-old congenitally blind student described the urban insects (cockroaches) in his home: the scratching sound, the damp smell, the segmented body under his fingers. He had never seen one, and his account came entirely from hearing, touch, and what others had told him. Asked how many legs a cockroach has, he said ten, attributing the figure to a plush toy he had handled for years.

That is not a defect in his account. Congenitally blind people build mental images from what their other senses give them, not from remembered sight, and those images hold together on their own terms (Cattaneo & Vecchi, 2011; Cattaneo et al., 2008). This study asks how such a student builds the shape of an animal he has never seen, whether generative tools can carry that shape forward instead of correcting it, and what a working process must do to keep him in the author's role.

Materials for blind students are built to deliver content. Braille, tactile graphics, and multimodal books all make it reachable (Lahav & Mioduser, 2004; Valente et al., 2024). Blind learners have been brought into design decisions, though rarely in schools. Three of 21 teachers of the visually impaired involved students in making the materials they would later use (Phutane et al., 2022). Participation at that stage shapes how this subject is presented. The subject is still chosen by teachers, publishers, and curriculum designers.

A route in the other direction runs through three technologies. Language processing can work from what the learner says and keep his sensory vocabulary instead of replacing it with a visual one. A text-to-image model turns that description into a picture. Swell-paper printing returns the picture to his hands. This study traces that sequence through a single sustained case: the co-creation of a 24-page tactile picture book with the student described above. The design brief was set by his sensory account rather than by an existing curriculum framework, and this paper reports what that reversal produced.

## Literature Review

### Sensory Perception and Learning Characteristics of VI Students

Blind and visually impaired learners come to rely more on hearing and touch than on any remaining sight (Cattaneo & Vecchi, 2011; Cattaneo et al., 2008). Touch recognizes an object by sampling it part by part. This is slower than vision, but what it builds is stable. Teaching methods adapted to non-visual channels help blind students learn about shapes and raise their motivation to learn (Wu, 2023).

The point that matters here is about images rather than about access to information. Congenitally blind people form and handle mental images built from what their senses give them, not rebuilt from remembered sight (Cattaneo et al., 2008). An image made this way is not a rough copy of a visual original. It is a picture with its own logic. Valente et al. (2024)

tested how well children with visual impairments understand a multimodal book, taking comprehension as the outcome. That question is about reception: how well a reader takes in content prepared by others. What happens when the blind reader is also the source of that content is a different question, and it is the one this study takes up.

## AI in Education and the Direction of Sensory Translation

Generative AI and natural language processing are already used in teaching, as adaptive tutors and as tools for making content (Holmes & Tuomi, 2022). In special education, AI has been used for visual impairment mainly in one direction: it turns visual material into non-visual output for a blind user (Khlaif et al., 2026). Automatic image description, scene narration, and AI-made tactile graphics from existing diagrams all run this way, as does the text-to-touch pipeline of Tanaka and Ochiai (2023).

In all of these the blind user is at the receiving end. The content starts with a sighted author, a curriculum, or a camera. AI translates it. The blind user takes in the result. This makes access the goal, and the design question becomes how much of the visual original survives the conversion.

The same technology can run the other way. Because it works from what the learner says, it can keep his sensory vocabulary instead of replacing it with a visual one. Paired with image generation, this makes his description the starting point of the translation rather than its destination, and each round of his comments changes what the system produces next.

## Tactile Design and Swell-Paper Production

A tactile image only works if the hand can follow it. That means clear outlines, a clear order to explore in, and shapes the reader can recognize (Wu et al., 2022). Thermal swell paper is the most common way to make such images. It is cheap and needs only heat unit by printing (Phutane et al., 2022). Good design also needs varied line weights, enough space between parts, Braille labels, and good to go with the touch (Wu et al., 2022).

Tactile picture books sit inside this literature. Valente et al. (2024) test how well children with visual impairments understand a multimodal book. Both look at books made for blind readers. Sighted designers and teachers chose the subjects, wrote the stories, and made the design decisions. Researchers have answered this by bringing blind learners into design, but little of it has reached classrooms. Phutane et al. (2022) interviewed 21 teachers of the visually impaired and ran a three-week diary study with eight of them. Three had involved students in making tactile materials. Where they did, the making itself became the lesson, and students used the results with more confidence. Participatory design is well established in assistive technology research and largely missing from education, and earlier systems either automated the making or treated the teacher as the only user. The same interviews show why substitutes are common. Teachers wanted three-dimensional models and real objects, could not get them, and made replacements from craft supplies and cheap household goods.

## The 4F Reflection Model

The 4F model sets out four stages of reflection: recording what happened, noting how it felt, working out what it means, and planning what to do next. It grows out of the experiential learning tradition running from Dewey through Kolb (2015) and Schön (1983), and cycles of

this kind give participants a sequence for turning experience into words. Here it works as a turn-taking structure rather than a measure. At each round the author had to say what he noticed, how he felt, what he concluded, and what he wanted changed, so the next image answered him rather than a designer.

### Research Gap

Each of these literatures stops at the same point. Research on blind cognition shows that congenitally blind learners build coherent images from non-visual input (Cattaneo et al., 2008), but does not ask what happens when those images are made available to others. Work on AI in visual impairment education has developed the visual-to-tactile pipeline in detail (Khlaif et al., 2026; Tanaka & Ochiai, 2023), but its direction is fixed: the source is visual, the blind user is the destination, and success means staying faithful to the original. Studies of tactile materials have brought blind learners into the making, yet the teacher still chooses the subject (Phutane et al., 2022). The gap is about position, not technology, since the tools themselves are available and documented (Lahav & Mioduser, 2004).

## Methodology

### Research Positioning

This is practice-based design research. What it produces is an object, a 24-page tactile picture book, and the knowledge it offers is the kind such research offers: an account of how something was made, the decisions taken along the way, and principles that carry over to similar work. It does not offer general measures of learning. The single participant is not a sample. He is the author of the book, and the study reports what became possible once he was given that role.

### Participant and Ethical Framework

One 15-year-old student, congenitally blind with no visual experience from birth, participated as the author. His guardians and school provided written consent, and his own assent was obtained at the outset and confirmed at each stage. He directed creative decisions, approved all outputs, and could decline to proceed at any point. Interviews were conducted in his home, where he could move and handle objects with confidence. Identifying details other than age and visual status have been withheld.

Two features of this arrangement were adopted deliberately. The student was named as author of the resulting book rather than described as a participant in it. Where his account of an animal conflicted with the biological record, the conflict was documented rather than resolved, and no attempt was made during production to bring his descriptions into line with the taught form.

### Tools

Google Labs Mixboard, an interface that combines language understanding with generative image editing, turned the author's spoken descriptions into images. He directed all revisions. Researchers worked the interface and turned his instructions into prompts. A 4F protocol shaped three formal interviews, covering sensory facts, feelings and creative ideas, and the building of the story. Final images were made into tactile graphics on a TactPlus thermal swell-paper printer, which raises treated paper under heat into readable relief. This is a raised-line method, not additive fabrication. No three-dimensional printing was used at any stage.

## Procedure

### *Stage 1 (Months 1–2): Interview and Exploration*

Three formal interviews followed the 4F structure. The author described household insects, the sensory experiences attached to them, and his emotional responses, and the team documented his language, metaphors, and comparisons. Preliminary character concepts emerged here.

### *Stage 2 (Months 2–5): Co-creation*

His descriptions were transcribed into prompts and submitted to Mixboard by an University student. He reviewed each generated image through a researcher's verbal description and issued revisions. Twenty-four cycles were completed across the two characters, a cockroach and a gecko, and the narrative developed alongside the images with his approval at each part.

### *Stage 3 (Months 5–8): Production and Author Verification*

Final images were converted to tactile form using the thermal swell-paper printer, and the 24-page book was completed with Chinese and English text alongside Chinese Braille. The author then read the finished pages by touch. He checked mainly the details of setting that carry the sensory core of each part, among them the textures for humidity, the stickiness of the trap, and the wall cracks, and he confirmed that the printed surfaces matched the scenes he had described. He expressed clear satisfaction with the finished object and with the story as his own. Reading by other visually impaired students is planned as the next stage of this work and had not been conducted at the time of writing.

## Data

The corpus comprises roughly fifteen hours of verbatim interview recordings, field notes on the author's reactions and creative decisions, an archive of more than 240 generated images across all iteration cycles, recordings of his responses to images and printed pages, process photographs, and the book itself.

# Findings

## The Author's Sensory World

Asked about the cockroaches in his home, the author did not describe their appearance. He described the scratching in the walls, the damp smell near the kitchen, and one remembered occasion of touching one: smooth segments, legs in sequence. His knowledge came from three sources of unequal weight. Most came from what family members had told him. A plush cockroach he had handled many times gave less, but what it gave was more specific. The direct encounter gave least of all, being brief and offering no chance to count or compare.

His language for movement was precise in ways visual description does not capture. The spoke of cockroaches rushing like water through cracks, and of the body trembling when touched. These temporal and textural terms carried into the images he later commissioned.

## The Sources of Structural Knowledge

Across the interview stages the author's accounts of animal form returned repeatedly to objects he had handled. Three cases are presented in descending order of evidential strength, since the objects differ in whether they could be examined.

### *The Frog: Object Available*

He described frogs as having forelimbs much longer than their hind limbs, which reverses the real proportion. The toy he named still exists and we examined it. It comes from the IKEA SAGOSKATT range, a limited yearly collection in which children's competition drawings are made into plush toys (IKEA Museum, n.d.). The published selection criteria cover originality, colour, expressiveness, and whether a design can be mass-produced. Anatomical accuracy is not one of them (IKEA, 2021). The toy sits in a crouch. Its forelimbs reach outward as slender continuous lines ending in separate digits. Its hind limbs fold against the body as a compact mass with less defined feet.

**Figure 1**
*The IKEA SAGOSKATT Frog*

Its proportions therefore come from a child's drawing that was made up without correction, which places the distortion before the toy rather than in it. Its posture also sets the forelimbs along a longer path than the hind limbs, so a hand following the shape measures the path it travels rather than the length of the limb.

Neither feature matters for a sighted child, who sees frogs in ponds, in books, and on film, and reads the toy as a stylized version. For this author, touch was the only way to know the toy, and the toy was his only reported source for the proportions of a frog. What a sighted audience reads as charm, and corrects without effort, reached him as the only account of the shape he had.

### *Two Further Cases, Objects Unavailable*

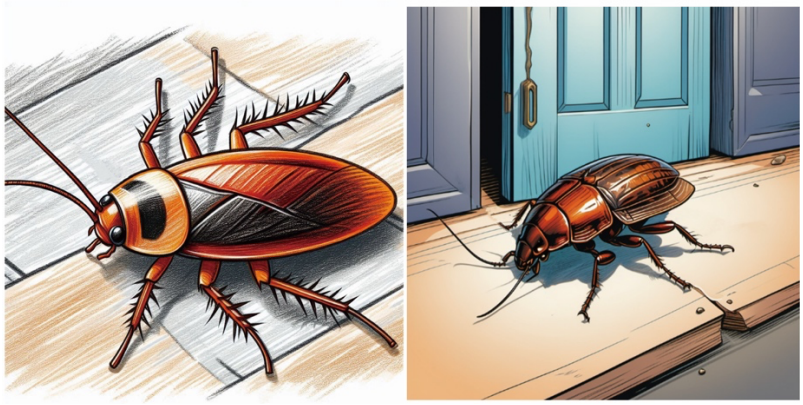
Asked how many legs a cockroach has, he answered ten, attributing this to a plush cockroach his family bought when he was around ten. The toy no longer survives, so the attribution is reported as his account rather than a verified property of the object. Two readings remain open: the toy may have been made with ten legs, or it may have had six alongside antennae, cerci, or seams indistinguishable from legs under the fingers, in which case the count arose in the reading. Asked separately whether a praying mantis has six legs, he again answered ten, again

citing a plush toy; none survives and the team holds no record of one. Insect anatomy is taught in his elementary curriculum and he had received that teaching. The mantis case is of interest for its scope rather than its object. Ten legs was not a fact he held about cockroaches but the body plan he held for insects, applied to a second species when he was asked.

## Directing the Image

The author's descriptions had to become pictures, and the first round showed what that risked. He had asked for a cockroach with ten legs. The model returned two images, each showing a conventional six-legged insect (Figure 2). The training data holds the standard body plan, and it supplied that plan instead of the description it was given. This was not a failure of prompting. It was the model imposing what it already held on an author who departed from it, and it marks the point where his version could most easily have been lost.

**Figure 2**
*First-Round Outputs for the Protagonist, Showing the Six-Legged Form Produced by the Model Before Revision*



The images were described to him aloud. He rejected the leg count himself, without prompting from the research team, and required ten. The revision was made, and the ten-legged form held through the twenty-four cycles that followed. The elder cockroach who appears later in the narrative was given ten legs on the same basis.

Later revisions followed the same pattern. Early antennae were too delicate, and he corrected the researchers directly: they needed to be stronger, with more feeling in them. A version was rejected because the legs were resting rather than poised to run. His instructions concerned posture and readiness rather than appearance, which is what a description assembled from sound and touch would be expected to priorities.

## World-Building Without Vision

Across the narrative the author specified details he cannot perceive. A wall calendar reads 1968; a second, in a later scene, reads 30 April 1981. Two Giant electric bicycles stand outside a lobby, one blue carrying the plate 1856 and one white carrying 93156. A yellow storage cabinet beside a sofa is priced at 1,111 yuan. Five violets sit by a window. The building manager reads a newspaper, on the left.

None of this reaches him through hearing or touch. Years on calendars, digits on plates, prices on labels, the position of a figure in a frame are all registered by sight, and he supplied them

without hesitation. Some of this is the conventional furniture of verbal description. He has spent years with audio recordings of Jin Yong's novels, where rooms contain calendars, calendars carry years, vehicles carry plates, and goods carry prices. He knows what a described world contains and populated his own accordingly. The values are arbitrary. Their presence is not.

### Building the Narrative

Before co-creation began, the research team taught the author the classical four-part Chinese narrative structure of qi, cheng, zhuan and he: opening, development, turn, and resolution. He then wrote to that structure. It came from the team, not from him, and we report it as a scaffold rather than as something he found for himself. What went inside it, and in what order, was his.

The 4F protocol served a different function, organizing the interviews rather than the story. The facts stage recorded what he had experienced: scratching in walls, water pipes, the damp smell near the kitchen. The feelings stage drew out his response, and he described the insects as trapped, searching for food, and frightened. In the findings stage this sensory knowledge met a larger idea, drawn from the martial arts fiction he listens to, of a protagonist facing impossible odds.

A cockroach wandering beneath a bedroom cabinet enters a newly laid trap out of curiosity and wakes to find others stuck around him. Relatives arrive, and an elder proposes sending the smallest out for help. The small one recruits geckos, who agree while privately anticipating a meal. They carry the trap to a puddle and free the colony, then turn on them, and fifty-eight are eaten. One escapes, crosses a pine wood, and reaches an apartment lobby, where it finds an abandoned nest behind a sofa. Eating crumbs there, it decides to open a school of cockroach survival, teaching others to evade geckos, avoid traps, and recognize human footsteps.

In the future stage we asked how he would explain the story to another blind student without the book in hand. He said he would describe what he had felt, the trapped feeling, then hope, then the shock of betrayal, because that was the real story. His written ending makes the same move. The protagonist is content at the close not because it survived, but because it has finally worked out how to live.

## Discussion

### AI as Translation Medium

The generative model worked neither as a teacher nor as a corrector but as a translator. It took the author's own sensory vocabulary as input and did not ask him for visual terms he did not have. Given descriptions full of tactile metaphor and timing, it produced images he could then revise. The tool moved from delivering information to carrying expression outward.

This depended on an arrangement that had to be kept up deliberately, and the first round showed why. Asked for the author's cockroach, the model produced a six-legged one, and the designers working the interface held the same expectation. Every point where his description differed from it was a point where the difference could have disappeared without anyone deciding to remove it. What kept the ten legs was not the tool and not the operators. It was the rule that only the author issued revisions, and the fact that he issued this one.

### The Position of the Doubters

Every challenge to the ten-legged form came from sighted people, and each drew on a channel the author cannot access: textbook illustrations, diagrams, remembered images. For those raising it, the objection that a cockroach has six legs is visually acquired. For the author it had only ever been a verbal report, and it did not displace what his hands had established.

The sighted responses divided. When design students at our institution first saw the generated character, one said it looked like an alien, then that it had the quality of a strange and marvellous creature, and that the author's world was unlike the one their own visual education had furnished. The strangeness registered first as error and then as interest. At a public presentation, an audience member asked whether the work had any meaning at all.

The answer is the position of this paper. As biology instruction the work has no value and could mislead. As creative work it opens a world that visual convention does not constrain. Both responses mark the same line from opposite sides, between judging a blind author's imagery against sighted norms and reading it on his terms. The ten-legged cockroach shows that this study held the second position. A finished character with six legs would have meant that the priors of the model and its sighted operators had replaced the author's account.

### What This Study Yields

The study yields an object, a documented process, and four principles that carry over to similar work. The first concerns input. The author's words should reach the model unedited, because turning them into visual vocabulary is where authorship is first lost. The second concerns control: only the author issues revisions, and collaborators work the interface without judging it. Third, places where his account differs from the biological record belong in the record of the process, not among its corrections, since they are what the process exists to carry. The fourth is checking. An object the author cannot read for himself has been made for him rather than by him.

### Limitations

This is a single case. The study reports what one working process made possible. It does not measure learning, retention, or transfer. The narrative structure was not the author's own discovery, since the four-part form was taught to him at the start and only the content inside it was his. The book has two characters, so nothing here speaks to scale. No other visually impaired student has read it yet, so nothing can be said about how it works for its audience. Of the three cases of knowledge about shape, only one involves an object we could examine. The interfaces were built for sighted operators, which is why researchers rather than the author worked them, and questions of ownership over co-created work remain open.

## Conclusion

This study set out to see what would follow from putting a congenitally blind student in the author's role and using generative tools to carry his account outward rather than correct it. The result is a 24-page tactile picture book whose main character has ten legs, made through a documented sequence that runs from spoken description to prompt, to generated image, to raised-line print, and checked by the author's own hands.

Three contributions follow. The paper documents a working method for authorship without vision and says where control over the work is gained and lost within it. It records cases in which an author's knowledge of animal shape came from handled objects never meant to teach, one of which can be traced back to how it was made, which suggests that the shortage of purpose-made tactile materials has effects reaching further than access. And it describes a kind of world-building that draws on learned narrative convention as well as on sensory record, visible in the calendars, plates, and price tags he specified for scenes he cannot see.

The ten legs carry the argument. The model itself gave six on the first attempt, drawing on the anatomy in its training data, and a finished cockroach with six legs would have meant that this expectation, shared by the sighted collaborators, had quietly replaced the author's account. The ten legs did not survive by default. The author heard the images described, rejected the count, and asked for them back. That he did so, and that they held through every later cycle, is the clearest sign available that the author's role, once given, was used.

## Acknowledgements

We thank the student author, whose work this book is, and his family and school for their support throughout. This work was supported by the National Science and Technology Council, Taiwan, under grant NSTC 114-2410-H-218-007-MY2 and NSTC 114-2813-C-218-005-H.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

Generative AI was used in this project in two distinct capacities, which are declared separately. As a research instrument, Google Labs Mixboard was used to generate and revise the images that constitute the artefact under study. This use is the subject of the research and is described in full in the Methodology. As a writing aid, the authors used Claude (Anthropic) to improve the language, clarity, and structural organization of the English text. No text was generated for substantive content, and no analysis, finding, or interpretation originated with the tool. The authors reviewed and edited all output and take full responsibility for the content of this publication.

## References

Cattaneo, Z., & Vecchi, T. (2011). *Blind vision: The neuroscience of visual impairment*. MIT Press.

Cattaneo, Z., Vecchi, T., Cornoldi, C., Mammarella, I., Bonino, D., Ricciardi, E., & Pietrini, P. (2008). Imagery and spatial processes in blindness and visual impairment. *Neuroscience and biobehavioral reviews*, *32*(8), 1346–1360. https://doi.org/10.1016/j.neubiorev.2008.05.002

Coulson, D., & Harvey, M. (2013). Scaffolding student reflection for experience-based learning: a framework. *Teaching in Higher Education*, *18*(4), 401–413. https://doi.org/10.1080/13562517.2012.752726

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, 57*(4), 542–570. https://doi.org/10.1111/ejed.12533

IKEA. (2021, March 9). *Five kids drawings come to life for SAGOSKATT collection*. IKEA Global Newsroom. https://www.ikea.com/global/en/newsroom/collaborations/five-kids-drawings-selected-to-become-the-soft-toy-limitededition-sagoskatt-collection-210309/

IKEA Museum. (n.d.). *SAGOSKATT soft toys from 2015*. Retrieved July 30, 2026, from https://ikeamuseum.com/en/explore/product-stories/sagoskatt-soft-toy-2016/

Khlaif, Z. N., Alshakhshir, R., Hamamra, B., & Joma, A. (2026). Assistive generative AI for visually impaired learners: Personalization and inclusion in higher education. *International Journal of Artificial Intelligence in Education*, *36*(1), 100008. https://doi.org/10.1016/j.ijaied.2026.100008

Kolb, D. A. (2015). *Experiential learning: Experience as the source of learning and development* (2nd ed.). Pearson Education.

Lahav, O., & Mioduser, D. (2004). Exploration of Unknown Spaces by People Who are Blind Using a Multi-Sensory Virtual Environment. *Journal of Special Education Technology*, *19*(3), 15–23. https://doi.org/10.1177/016264340401900302

Phutane, M., Wright, J., Castro, B. V., Shi, L., Stern, S. R., Lawson, H. M., & Azenkot, S. (2022). Tactile materials in practice: Understanding the experiences of teachers of the visually impaired. *ACM Transactions on Accessible Computing, 15*(3), 1-34. https://doi.org/10.1145/3508364

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Tanaka, K., & Ochiai, Y. (2023). From text to touch: Methods and case studies for designing tactile graphics via digital fabrication and AI. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 1–13). ACM.

Valente, D., Chennaz, L., Archambault, D., Négrerie, S., Blain, S., Galiano, A. R., & Gentaz, E. (2024). Comprehension of a multimodal book by children with visual impairments. *British Journal of Visual Impairment*, *42*(1), 276–286.

Wu, C. F., Wu, H. P., Tu, Y. H., Yeh, I. T., & Chang, C. T. (2022). A study on the design procedure of three-dimensional printable tactile graphics for individuals with visual impairments. *Journal of Visual Impairment & Blindness, 116*(4), 507–516. https://doi.org/10.1177/0145482X221122754

Wu, H. P. (2023). Improving learning of figurative concepts in individuals with blindness: Adopting teaching strategies to enhance learning motivation. *Journal of Visual Impairment & Blindness, 117*(4), 292–302. https://doi.org/10.1177/0145482X231193021

# Reflective Practice in Pre-professional Arts Education: A Theoretical Examination Through Teacher Education

Misun Kong, Konkuk University, South Korea
Minsung Kwon, Chung-ang University, South Korea
Jonghun Kim, Konkuk University, South Korea

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This study explores how reflective practice theories developed in teacher education can be conceptually reconstructed for pre-professional arts education in the Korean higher education context. Rather than assuming that teachers and artists are similar professional groups, the study focuses on their shared educational position as pre-professionals who develop professional practice within institutional education. In Korea, teacher education prepares pre-service teachers for professional teaching practice, while arts colleges place strong emphasis on studio, performance, and creative practice that prepares students for professional work in their artistic fields. This study conducts a theoretical and conceptual analysis of Korthagen's ALACT reflective cycle and Darling-Hammond's practice-based teacher education, with Schön's reflective practitioner theory as a supporting perspective. The analysis identifies a shared learning structure involving authentic practice, reflection on experience, awareness of important aspects, exploration of alternatives, and renewed practice. However, this structure should not be directly transferred to arts education. Instead, it needs to be reconstructed in relation to the open-ended and context-dependent nature of creative practice. The study therefore proposes a reflective cycle of creative practice, reflection, awareness, alternative exploration, and re-practice. It also identifies documentation, dialogue, feedback, and opportunities for renewed practice as important educational conditions. The study provides a theoretical basis for reconsidering professional learning in Korean higher arts education through reflective practice.

*Keywords:* reflective practice, pre-professional artists, teacher education, professional learning, arts education

iafor
The International Academic Forum
www.iafor.org

## Introduction

Practical and creative experience is an important part of professional development in arts education. Students majoring in the arts repeatedly engage in activities such as performance, acting, dance, visual production, filmmaking, and other forms of creative work. Through these experiences, they develop technical skills and gradually build their own ways of artistic expression. In the Korean higher education context, arts colleges generally place considerable emphasis on practice within each major and prepare students for professional practice in their artistic fields. In this sense, creative practice in higher arts education can be understood not only as a learning activity but also as part of the transition toward professional artistic work.

However, having more creative experience does not necessarily mean that professional learning has occurred. Repeated practice may improve technical skills, but repeating the same action does not automatically help learners understand their own decisions or adjust them in new situations. Schön (1983) argues that professionals are not simply people who apply established knowledge and rules. They work in uncertain situations, interpret problems, make judgments, and adjust their actions through practice. Eraut (1994) similarly explains that professional knowledge and competence develop through practice, but experience needs to be interpreted and connected with existing knowledge.

From this perspective, reflection can be understood as a process that helps transform repeated practice into professional learning. Dewey (1938) argued that not all experience is educational. Experience becomes educational when it is connected with previous and future experiences and is reconstructed through interaction with the environment. This argument is also relevant to arts education. Providing students with many creative activities may not be sufficient if those experiences remain separate from one another and are not examined in relation to later practice.

Previous research has applied this perspective to arts education by arguing that experience becomes educational not simply because an activity takes place, but because experience is organized through continuity, interaction, and processes that allow learners to make meaning from it (Kong & Kwon, 2026b). From this perspective, the educational question is not only what kind of experience should be provided, but also how one experience can be connected to and transform later experiences.

This issue has become more visible with the rapid development of generative artificial intelligence. Generative AI can produce images, music, videos, and text at increasing speed and quality. This development makes it necessary to reconsider the relationship between the final creative product and the process through which the product is created. Kong and Kwon (2026a) addressed this issue by proposing an AI-mediated extracurricular program that organized artistic learning through repeated cycles of action, reflection, and re-practice. In that study, generative AI was not understood as a substitute for the creator but as a mediator that could help document, externalize, and revisit the creative process.

Such program-level discussions, however, raise a more fundamental theoretical question: Why can reflective practice theories developed in teacher education be considered relevant to the education of future artists?

The present study does not answer this question by assuming that teaching and artistic practice are the same. Teachers engage in educational practice, while artists engage in creative practice. Their professional purposes, working conditions, and criteria for judgment are clearly different.

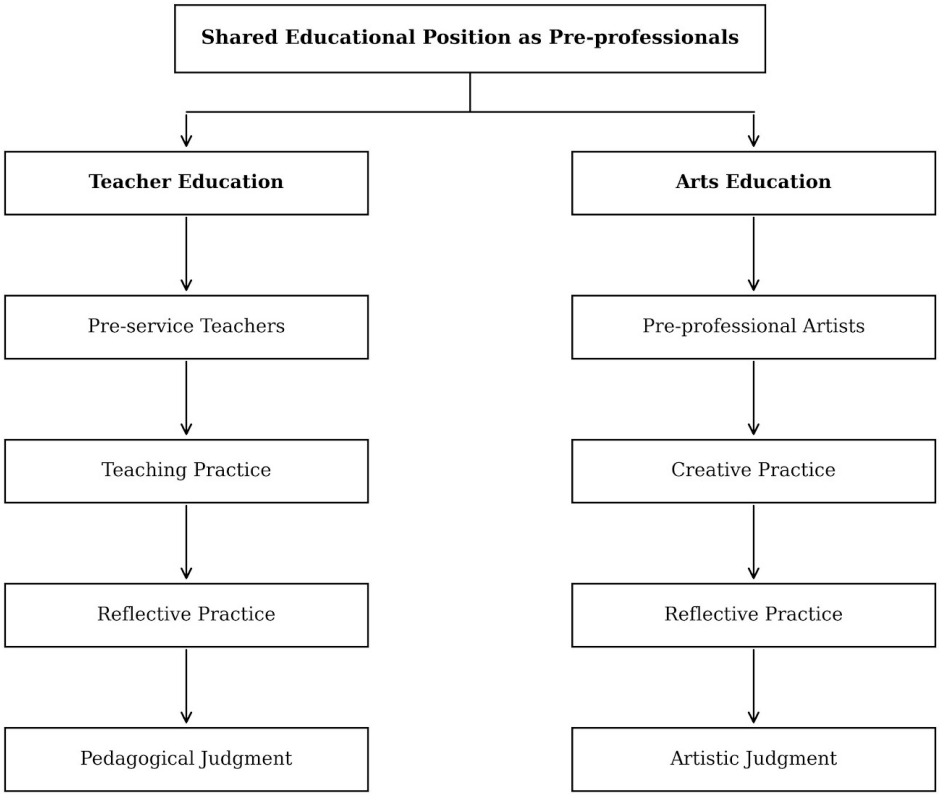
Therefore, directly transferring teacher education methods to arts education would overlook important differences between the two professional fields.

Instead, this study focuses on the educational position of learners before entering professional practice, particularly in the Korean higher education context. Students in colleges of education and teacher preparation programs are educated as pre-service teachers who are preparing to perform professional teaching practice. Similarly, students in arts colleges engage extensively in studio work, rehearsal, production, and performance as part of preparation for professional artistic practice. In this study, such students are described as pre-professional artists.

Pre-service teachers and pre-professional artists prepare for different professions, but both groups learn within institutional higher education before fully entering professional practice. They are expected to transform knowledge and experience into professional judgment through actual practice. This shared position as developing professionals provides a basis for examining whether theories of reflective professional learning developed in teacher education may also provide useful theoretical resources for arts education.

**Figure 1**

*Structural Connection Between Teacher Education and Pre-professional Arts Education in the Korean Higher Education Context*




The purpose of this study is therefore to examine the main principles of Korthagen's reflective cycle and Darling-Hammond's practice-based teacher education and to conceptually reconstruct these principles for pre-professional arts education in Korea. The central research question is: How can reflective practice theories developed in teacher education be reconstructed for pre-professional arts education in the Korean higher education context?

## Theoretical Background

### Reflective Practice and Professional Learning

The concept of reflective practice challenges the idea that professional expertise is simply the accumulation of knowledge and technical skills. Schön (1983) argues that professionals often work in situations where problems are unclear and where existing rules cannot provide complete answers. Professional practice therefore requires continuous interpretation, judgment, and adjustment.

Schön distinguishes between reflection-in-action and reflection-on-action. Reflection-in-action occurs while a person is engaged in practice. The practitioner recognizes a problem or unexpected situation and adjusts an action while it is taking place. Reflection-on-action occurs after the practice has taken place and involves reconsidering the experience, the decisions that were made, and their consequences.

Reflection in this sense is different from simply identifying mistakes after completing a task. It involves examining experience, understanding the decisions behind an action, and using that understanding to inform later practice. Schön's work is particularly relevant to the present study because his discussion of reflective professional practice is not limited to teaching. His analysis includes fields such as architecture and design, where practitioners make judgments in situations of uncertainty, uniqueness, and complexity.

Moon (2004) also describes reflection as a learning process through which individuals reconsider experience and develop new understanding. Experience and reflection are therefore not separate parts of learning. Reflection connects one experience with another and enables the learner to reinterpret earlier actions.

This perspective is relevant to creative practice. Artists continually make choices in situations where there is no single correct answer. They choose materials, forms, movements, sounds, images, and ways of presenting ideas. Simply repeating a technique does not necessarily prepare a learner to respond to new artistic problems. Pre-professional artists therefore need opportunities to examine why they made particular choices, what those choices produced, and what other approaches might have been possible.

### Korthagen's ALACT Reflective Cycle

Korthagen et al. (2001) describe the ALACT model as a reflective cycle that connects practice and reflection in teacher education. The model consists of five stages: Action, Looking back on the action, Awareness of essential aspects, Creating alternative methods of action, and Trial.

The first stage, Action, refers to actual practice. The learner then looks back on the experience and considers what happened. In the Awareness stage, the learner identifies important aspects of the experience, including patterns, assumptions, or factors that influenced the action. The learner then considers alternative approaches and finally tries a new form of action.

Trial becomes a new Action, meaning that the model does not end after one cycle. Reflection becomes part of an ongoing process of practice, reconsideration, and renewed practice.

A key feature of the ALACT model is that reflection does not end with evaluation. The learner does not simply ask whether an action was successful or unsuccessful. Reflection should lead to awareness of why a particular action occurred and to consideration of what could be done differently in the future. Korthagen and Vasalos (2005) further argue that deeper reflection may move beyond observable behavior and competence to include beliefs, identity, and professional meaning.

This cyclical structure has a possible connection with creative practice. Pre-professional artists also act, review what they have done, recognize aspects of their creative choices, explore other possibilities, and try again. However, this similarity does not mean that ALACT can be directly transferred from teacher education to arts education.

Teaching and creative practice operate according to different professional purposes and forms of judgment. Artistic practice is particularly open-ended, and different artistic choices can have different but equally meaningful outcomes. The present study therefore focuses not on transferring the specific methods of ALACT, but on reconstructing its underlying principles of practice, reflection, awareness, alternative exploration, and renewed practice for creative learning.

**Darling-Hammond's Practice-Based Teacher Education**

Darling-Hammond (2012) argues that professional teacher education should move beyond a fragmented structure in which theory is learned first and applied later. Instead, professional knowledge should develop through the integration of theory, authentic practice, analysis, and reflection.

From this perspective, pre-service teachers learn not only by studying theories of teaching but also by experiencing real or realistic teaching situations, examining problems that arise in practice, and connecting those experiences with theoretical ideas. Professional learning therefore takes place within practice rather than after practice.

Practice-based professional education emphasizes authentic learning experiences, analysis of practice, feedback, and collaboration. Darling-Hammond et al. (2017) similarly identify active learning, collaboration, feedback, reflection, and learning connected with actual professional problems as important features of professional development.

Korthagen and Darling-Hammond can therefore be understood as addressing different but related dimensions of professional education. Korthagen explains how a learner moves from one experience through reflection toward new action. Darling-Hammond provides a broader explanation of the educational environment in which such professional learning can occur.

Taken together, the two perspectives suggest three principles. First, authentic practice is the starting point of professional learning. Second, experience needs to be analyzed and reconstructed through reflection. Third, reflection should lead to alternative actions and renewed practice. Repeated cycles of this process can support the gradual development of professional judgment and agency.

## Method

This study is a theoretical and conceptual study rather than an empirical investigation. It does not analyze interviews, observations, surveys, or learner-generated data. Its purpose is to examine the conceptual structure of reflective practice theories developed in teacher education and to reinterpret those ideas within the context of pre-professional arts education in Korea.

The analysis was conducted in three stages.

First, the study examined Korthagen et al.'s (2001) ALACT reflective cycle and Darling-Hammond's (2012) practice-based teacher education to identify how each perspective explains the relationships among practice, reflection, and professional learning. Schön's (1983) concept of the reflective practitioner, Eraut's (1994) discussion of professional knowledge and competence, and Moon's (2004) work on reflective and experiential learning were used as supporting theoretical perspectives to clarify the broader relationship between experience, reflection, and professional development.

Second, the educational position of pre-service teachers was compared conceptually with the position of students preparing for professional artistic practice in Korean higher arts education. The comparison did not focus on occupational similarity. Instead, it focused on the structural position of both groups as learners who engage in authentic practice within institutional education before fully entering professional work.

Third, the principles identified from teacher education were reconsidered in relation to the specific characteristics of creative practice. Elements that could be conceptually shared and elements that required reinterpretation were distinguished. Based on this analysis, a reconstructed structure of reflective practice for pre-professional arts education was proposed.

The arts education example presented later in the paper is therefore an illustrative example, not empirical evidence collected from an actual course or group of students.

## Conceptual Findings: Reconstructing Reflective Practice for Pre-professional Arts Education

### Pre-professionals as a Shared Educational Position

The conceptual analysis suggests that the theoretical connection between teacher education and arts education is better understood through the educational position of their learners than through similarity between the professions themselves.

In the Korean context, pre-service teachers study educational and subject-specific knowledge while also participating in activities such as lesson planning, microteaching, and teaching practice. These experiences are designed to prepare learners to make professional judgments in actual educational situations.

Students in arts colleges similarly develop knowledge and skills through rehearsal, studio work, production, performance, and other forms of creative activity. Although the balance between theoretical and practical education varies across disciplines and institutions, practical creative work is an important part of professional preparation in higher arts education.

This comparison does not suggest that teachers and artists perform the same kind of professional work. Rather, both groups are learners preparing for different forms of professional practice.

This shared educational position provides the main theoretical reason for examining reflective practice from teacher education within arts education. Teacher education offers an established body of theoretical work on how learners move from institutional study toward professional practice through experience and reflection. The question is whether the underlying learning principles can be reconstructed for artistic professional preparation.

## From Practice to Reflective Re-practice

The conceptual analysis also identifies a cyclical structure of professional learning. When the perspectives of Korthagen and Darling-Hammond are considered together, professional learning can be understood as a cyclical process of practice, reflection, awareness, exploration of alternatives, and renewed practice. In this process, practice is not the end of learning but the starting point of further reflection. Learners revisit their experience, identify important aspects of their decisions and circumstances, consider alternative approaches, and return to practice with a revised understanding.

This distinction is important because repetition and reflective repetition are not the same. Repeating the same artistic exercise may improve speed, coordination, or technical accuracy. However, repetition alone does not necessarily produce deeper professional judgment.

Reflection enables learners to examine relationships between experiences. It allows them to consider why they made certain choices, what assumptions shaped those choices, and whether a different approach might produce another result. In this sense, reflection changes repetition into a learning process.

This argument is consistent with Dewey's (1938) principle of continuity. Experience becomes educational when it shapes later experience rather than remaining an isolated event. Previous work in arts education has similarly argued that educational experience depends not simply on the number of activities students perform but on how experiences are connected and reconstructed (Kong & Kwon, 2026b).

## Conceptual Reconstruction for Creative Practice

A further point emerging from the conceptual analysis is the distinction between application and reconstruction.

Direct application would mean transferring the methods of teacher education into arts education without sufficiently considering the characteristics of creative practice. This study does not propose such an approach.

Teaching practice and creative practice involve different purposes and different criteria of judgment. In creative practice, decisions can be shaped by personal artistic values, aesthetic intentions, genre conventions, media, material conditions, and individual interpretation. There is often no single correct solution.

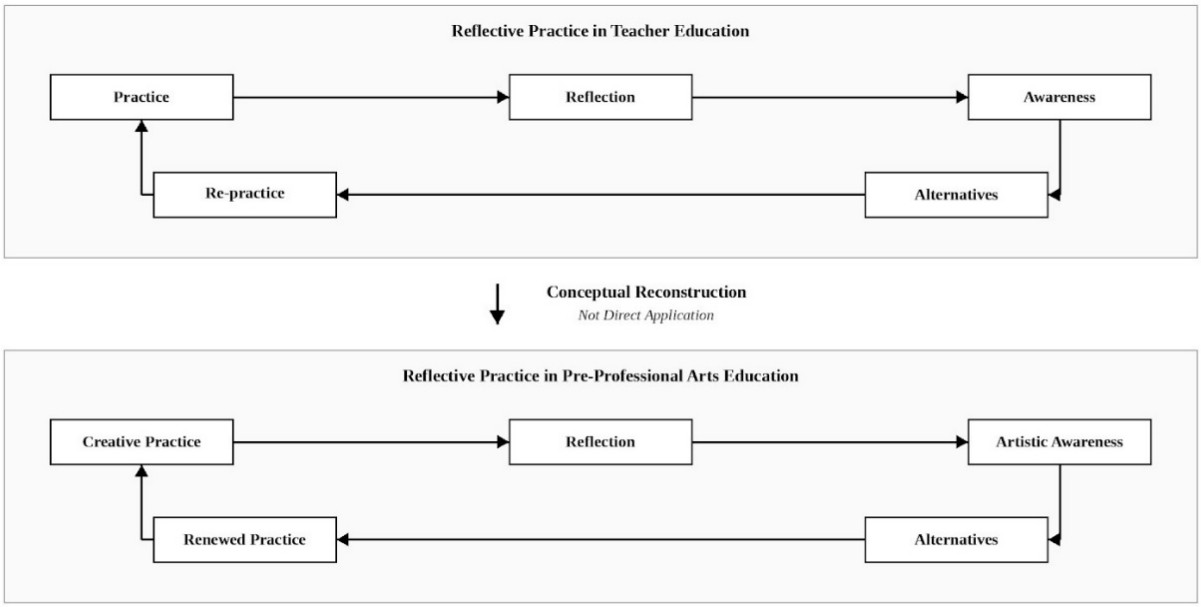
For this reason, the central principles of reflective practice need to be reconstructed in the language and context of artistic creation rather than transferred directly from teacher education. This approach is consistent with previous curriculum research emphasizing the reinterpretation of theoretical principles in relation to a new educational context. Kwon and Kong (2026), for example, reconstructed Dewey's concept of experience, Tyler's curriculum structure, and Schwab's practical perspective into an integrated experience–structure–context framework for film-based liberal arts education. Building on this approach, the present study reconstructs reflective practice for pre-professional arts education as a cyclical process in which learners engage in creative practice, reflect on their creative process and choices, become aware of recurring artistic patterns and judgments, explore alternative possibilities, and return to creative practice with a revised understanding.

Through repeated cycles, learners may begin to understand not only the outcomes of their creative work, but also how those outcomes are connected to the artistic choices they make.

This process should not be interpreted as a claim that reflection directly "produces" an artist's professional identity. Professional identity develops over time through multiple personal, social, institutional, and professional experiences. Reflective practice is better understood as one educational process that can support professional self-understanding, artistic judgment, and agency, which may contribute to longer-term professional formation.

**Figure 2**

*Conceptual Reconstruction of Reflective Practice for Pre-professional Arts Education*




## An Illustrative Example in Arts Education

A simple example from performing arts education can illustrate how this reconstructed process might operate.

Consider a student preparing a monologue. The student repeatedly rehearses vocal expression, movement, timing, and emotional interpretation. In a primarily outcome-oriented approach, the quality of the final performance and the instructor's evaluation may become the main indicators of learning.

From a reflective practice perspective, however, learning continues beyond the completed performance.

The student may review a rehearsal video and record why a particular movement, rhythm, or vocal choice was made. Through discussion with peers, the student may discover a difference between the intended expression and the way others understood the performance. The student may also notice a repeated habit, such as using similar movements whenever a particular emotion is expressed.

In the next rehearsal, the learner can try a different use of space, movement, pace, or vocal rhythm. The new performance can then be compared with the previous one to consider how different choices relate to the artistic intention.

The purpose is not to identify one correct performance. Rather, the educational focus is on the process of acting, revisiting choices, recognizing patterns, exploring alternatives, and returning to practice.

This example should not be understood as empirical evidence for the effectiveness of the model. It is an illustration of how the principles of reflective practice may be reconstructed while maintaining the open-ended character of artistic creation.

## Educational Conditions for Reflective Practice

Reflection does not automatically occur simply because students are asked to reflect on their work. Several educational conditions are needed if reflective practice is to become part of professional learning.

First, learners need authentic creative practice. Reflection requires an actual experience that can become an object of examination. For pre-professional artists, this may involve rehearsal, studio work, production, performance, or other forms of creative activity.

Second, learners need ways to document and revisit the creative process. Many decisions, revisions, failures, and changes disappear from the final product. Creative journals, rehearsal videos, audio recordings, sketches, portfolios, and other forms of documentation can make aspects of the process visible again. Kong and Kwon (2026a), for example, proposed documentation as a way to help learners externalize their creative process and examine patterns that might otherwise remain unnoticed.

Third, reflection needs dialogue and feedback. Individual reflection is important, but learners may continue to treat their own artistic habits as natural or unavoidable if they do not encounter other perspectives. Questions from peers and instructors can help learners reconsider their choices and recognize that an existing approach is one possibility among several.

Fourth, learners need opportunities to try again. Reflection that ends with a written evaluation or discussion may remain separate from professional practice. Learners need opportunities to turn the outcomes of reflection into new creative action.

These conditions also help explain why extracurricular learning may sometimes provide useful support for reflective practice. This does not mean that extracurricular learning is better than the formal curriculum or should replace it. Formal curricula require clear educational aims,

assessment criteria, and learning outcomes. However, environments with relatively lower assessment pressure may provide additional space for experimentation, failure, revision, and process documentation.

Generative AI may also be considered at this level. AI is not the central focus of the present study and does not replace human reflection. Rather, it may function as a reflective mediator by helping learners organize records, compare different stages of work, or revisit descriptions of their own creative process. Previous research has similarly conceptualized generative AI not as an agent that replaces creative judgment but as an environment that can support documentation and reflection (Kong & Kwon, 2026a).

Accordingly, extracurricular learning and generative AI should be understood as possible educational conditions or mediators that can support the cycle of practice, reflection, and re-practice rather than as the theoretical foundation of reflective arts education.

## Discussion

This study examined how reflective practice theories developed in teacher education can be reconsidered in pre-professional arts education. The analysis suggests that the connection between the two fields is not based on occupational similarity. Instead, it is based on the structure of higher education that prepares learners for professional practice.

This point is especially important for explaining why teacher education was selected as the theoretical starting point of the study. Teacher education has developed substantial theoretical work on how pre-service teachers experience professional practice, interpret their experiences, and gradually develop professional judgment. Korthagen and Darling-Hammond both address this transition from learner to professional practitioner.

Within the Korean context, higher arts education also has a strong practical dimension. Although curriculum structures vary across artistic disciplines and institutions, students commonly engage in sustained creative and performance-based practice as preparation for professional artistic work. For this reason, arts students can be understood as pre-professionals who are developing the ability to transform technical knowledge and creative experience into artistic judgment.

The two educational contexts therefore share a structural problem: professional practice cannot be learned through theoretical knowledge alone, but practice alone is also insufficient.

Learners need experience, but they also need opportunities to understand that experience. They need to recognize the judgments behind their actions, consider alternatives, and return to practice with a different understanding. This is the central point at which reflective practice from teacher education becomes theoretically relevant to arts education.

At the same time, the analysis also shows why conceptual reconstruction is necessary. Teacher education and arts education cannot share the same criteria for successful professional practice. Teaching involves educational aims, student learning, and pedagogical appropriateness, while creative practice involves artistic intentions, aesthetic judgment, interpretation, and the characteristics of particular media.

The reflective questions therefore need to change when the theory moves from one professional field to another.

The contribution of the present study can also be distinguished from earlier program design research. Kong and Kwon (2026a) used Korthagen and Darling-Hammond to propose an AI-mediated extracurricular program in which creative experience, documentation, reflection, feedback, and re-practice were systematically connected. That study primarily addressed the question of how reflection could be organized within a program.

The present study addresses an earlier theoretical question: Why is it reasonable to draw on teacher education theory when examining reflective learning among pre-professional artists?

The answer proposed here is the shared educational position of pre-professionals. Teacher education is therefore not treated as a model to be copied. It serves as a theoretical resource for understanding how practice becomes professional learning through reflection.

This distinction is central to the argument of the study. The value of reflective practice for arts education does not lie in transferring a teacher education model into another discipline. It lies in reconsidering the relationship among creative experience, reflection, and renewed practice within the professional preparation of future artists.

## Conclusion

This study examined reflective practice theories developed in teacher education and explored how their main principles can be conceptually reconstructed for pre-professional arts education in Korea. The analysis focused primarily on Korthagen's ALACT reflective cycle and Darling-Hammond's practice-based teacher education.

Four main conclusions can be drawn.

First, pre-service teachers and pre-professional artists prepare for different professions, but within the Korean higher education context they occupy a similar educational position as learners preparing for professional practice. This shared position provides a theoretical basis for reconsidering reflective practice across the two areas of professional education.

Second, the synthesis of Korthagen's and Darling-Hammond's perspectives identifies a common structure of professional learning involving practice, reflection, awareness, exploration of alternatives, and re-practice. Repeated practice alone does not guarantee professional development. Reflection helps learners interpret experience and connect it to later action.

Third, this structure needs to be conceptually reconstructed rather than directly applied to arts education. The study therefore proposes a cycle of creative practice, reflection on creative choices and processes, awareness of artistic judgment, exploration of alternative possibilities, and renewed creative practice.

Fourth, reflective practice requires educational conditions that support it. Authentic creative experience, documentation, interaction with peers and experts, and opportunities for renewed practice are important. Extracurricular learning and generative AI may support some of these conditions, but neither should be understood as a replacement for reflection itself.

The contribution of this study is therefore not the direct application of teacher education theory to arts education. Instead, it identifies a structural connection between two forms of pre-professional education and reconstructs the principles of reflective practice for the particular characteristics of creative learning. This perspective allows creative practice to be understood not only as technical training or production of artistic outcomes, but also as an educational process through which learners can gradually develop artistic judgment, agency, and professional self-understanding.

The study also has several limitations. First, it is a theoretical and conceptual study and does not include empirical data from pre-professional artists. The arts education example presented in the paper is illustrative rather than empirical. Future research should therefore examine actual reflective processes through materials such as creative journals, interviews, rehearsal records, peer feedback, and portfolios.

Second, the study is grounded primarily in the context of Korean higher arts education. The argument assumes a context in which arts colleges place substantial emphasis on studio, performance, and creative practice and in which higher education can serve as preparation for later professional artistic work. The purposes and curriculum structures of higher arts education, however, may differ across countries. For this reason, the structural relationship proposed between pre-service teachers and pre-professional artists should not automatically be generalized to other national contexts.

Future comparative research could examine how arts colleges in different countries understand professional preparation and whether the educational position of the “pre-professional artist” is shared across different higher education systems. Research across different artistic fields, such as music, visual arts, performing arts, and film, could also clarify how the objects and forms of reflection change according to the characteristics of creative practice.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, the author used ChatGPT (OpenAI) to support English translation, language editing, and the organization of written expression. Google Gemini was also used to generate the visual representations of the conceptual frameworks presented in Figures 1 and 2. The author reviewed and revised all AI-assisted outputs and takes full responsibility for the content of the manuscript.

## References

Darling-Hammond, L. (2012). *Powerful teacher education: Lessons from exemplary programs*. Jossey-Bass.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

Dewey, J. (1938). *Experience and education*. Macmillan.

Eraut, M. (1994). *Developing professional knowledge and competence*. Falmer Press.

Kong, M., & Kwon, M. (2026a). *Jeonmundaehak yesul jeongong haksaeng-ui changjak "gwajeong" seongchal gyoyuk-eul wihan AI maegae bigyogwa peurogeuraem seolgye yeongu* [Designing an AI-mediated extracurricular program to support reflective practice in art students' creative processes at junior colleges]. *Journal of Contents and Industry, 8*(1), 101–108. https://doi.org/10.55387/JCI.2025.08.01.15

Kong, M., & Kwon, M. (2026b). *"Gyeongheom"-ui yesulgyoyukjeok jeonhwan jogeon tamsaek: John Dewey gyeongheom gaenyeom-ui jaehaeseok-gwa yeonghwa giban gyoyang sueop bunseok* [A study on the conditions for the transformation of "experience" into arts education: Reinterpreting Dewey's concept of experience and analyzing film-based liberal arts classes]. *The Korean Journal of Arts Education, 24*(2), 23–34. https://doi.org/10.23317/kaae.2026.24.2.002

Korthagen, F. A. J., Kessels, J., Koster, B., Lagerwerf, B., & Wubbels, T. (2001). *Linking practice and theory: The pedagogy of realistic teacher education.* Lawrence Erlbaum Associates.

Korthagen, F. A. J., & Vasalos, A. (2005). Levels in reflection: Core reflection as a means to enhance professional growth. *Teachers and Teaching: Theory and Practice, 11*(1), 47–71.

Kwon, M., & Kong, M. (2026). *Yeonghwa kontencheu giban gyoyanggyoyuk-ui gyeongheom jungsim gyoyukgwajeong-e daehan ironjeok gochal: Dewey, Tyler, Schwab iron-ui jaeguseong-eul jungsim-euro* [A theoretical exploration of experience-centered curriculum in film-based liberal arts education: Focusing on the reconstruction of Dewey, Tyler, and Schwab's theories]. *The Korean Journal of Arts Education, 24*(2), 413–430. https://doi.org/10.23317/kaae.2026.24.2.023

Moon, J. A. (2004). *A handbook of reflective and experiential learning: Theory and practice*. RoutledgeFalmer.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

**Contact email:** kongms92@gmail.com

# Reconstructing Lexical Networks of Japanese EFL Learners: An Application of Technique Feature Analysis to Already-Known Words

Noriko Aotani, Tokai Gakuen University, Japan
Shin'ya Takahashi, Tokai Gakuen University, Japan

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

This study employed Technique Feature Analysis (TFA) criteria to analyze the characteristics of tasks effective for reconstructing Japanese EFL learners' lexical networks. We focused on the deepening and expansion of knowledge of already-known words. Twenty words were selected from an essay, and participants (73 Japanese university students) evaluated the perceived relationships among these words (Test 1). One week later, they read the essay and completed one of three tasks. Participants in the MCQ group answered ten multiple-choice questions regarding the meanings of the target words. The FIB group filled in ten blanks in the essay using target words provided with their meanings and usage examples. The Comp group completed ten sentences unrelated to the essay with the target words. TFA points assigned to the tasks were calculated as 5 for MCQ, 7 for FIB, and 8 for Comp. After completing the task, participants took the same word-relationship test (Test 2). Finally, four weeks after Test 2, participants completed Test 3. The results showed that word-relationship scores increased from Test 1 to Test 2 and decreased from Test 2 to Test 3 in all groups. No significant effects of group (i.e., task type) were found. The results were also analyzed and visualized using Gephi, a data visualization platform, revealing qualitative features of individual changes in lexical networks. The present findings did not correspond to the learning effects predicted by TFA, suggesting that more appropriate tasks and/or evaluation criteria are necessary for applying the TFA framework to the learning of already-known words.

*Keywords*: Japanese learners of English, lexical network, extensive reading, already-known words, Technique Feature Analysis

iafor
The International Academic Forum
www.iafor.org

## Introduction

Considerable debate has been devoted to the extent to which extensive reading contributes to vocabulary acquisition. Consistent with his Input Hypothesis, Krashen (1989) argued that extensive reading alone can lead to substantial vocabulary acquisition, and numerous studies have since attempted to examine this claim using a variety of methodological approaches. While most of this research has focused on the acquisition of new vocabulary, Horst, et al. (1998) suggested that, although extensive reading may play only a limited role in vocabulary acquisition for beginning second language learners, it nevertheless has the potential to deepen learners' knowledge of already-known words and to strengthen the lexical network by establishing new connections among these words. Pigada and Schmitt (2006) further argued that the effects of extensive reading on vocabulary learning should be evaluated from multiple perspectives, including form, meaning, and use, as represented in Nation's (2001) framework of nine aspects of word knowledge. Likewise, Waring and Takaki (2003) maintained that authors of graded readers should pay less attention to the introduction of new vocabulary alone and instead provide learners with richer collocational information for words they have already learned.

The conceptual model of the mental lexicon as a network structure has been widely accepted (e.g., Aitchison, 1994). This model is inspired by the fact that the human brain itself consists of interconnected neural networks, in which lexical items (nodes) metaphorically correspond to neurons and the links among them correspond to synaptic connections. An increase in vocabulary size is thus conceptualized as an increase in the number of nodes, whereas the enrichment of knowledge about an individual word is viewed as the creation of additional links. This network metaphor provides a compelling representation of how lexical knowledge may be organized in the human mind. Accordingly, a number of pedagogical studies have attempted to facilitate vocabulary learning by presenting semantically related groups of words. However, in most cases, these semantic relationships are theoretically determined rather than empirically derived from evidence showing how actual language users associate words in their mental lexicon (Wilks & Meara, 2002).

The Levels of Processing theory proposed by Craik and Lockhart (1972) argues that the depth of cognitive processing directly influences memory performance. Building upon this framework, Laufer and Hulstijn (2001) proposed the Involvement Load Hypothesis (ILH), which provides an index for predicting the effectiveness of incidental vocabulary learning induced by tasks accompanying extensive reading. Although numerous studies have examined the validity of the ILH, the vast majority have focused primarily on the retention of newly learned words. Motivated by this line of research, we conducted research which aimed to investigate how extensive reading, together with task-induced cognitive involvement as conceptualized in the ILH, promotes the deepening and broadening of lexical network structures centered on already-known words among Japanese learners of English.

This series of studies yielded two major findings (Aotani & Takahashi, 2022, 2023, 2025). First, tasks that are effective in promoting the retention of newly learned vocabulary are not necessarily effective in deepening or broadening the lexical network structures of already-known words. Second, the three components of psychological involvement proposed by the ILH—*need*, *search*, and *evaluation*—are insufficient for identifying the characteristics of tasks that facilitate lexical network restructuring. These findings suggest the necessity of adopting an alternative analytical framework for task evaluation.

Alongside the ILH, Technique Feature Analysis (TFA), proposed by Nation and Webb (2011), provides another influential framework for analyzing the characteristics of vocabulary learning tasks. Several studies have compared the validity of the ILH and the TFA, identifying both their strengths and limitations (Gohar et al., 2018; Hu & Nassaji, 2016; Qin, 2023). However, these studies have also focused exclusively on the acquisition and retention of newly learned vocabulary. The present study aimed to apply the TFA framework to examine the task characteristics that promote the restructuring of lexical networks, which could not be adequately explained by the ILH.

## Methods

### Participants

Seventy-three Japanese university students (45 males and 28 females, mean age: 20.5 years old) participated in the experiment as paid volunteers.

### Materials

An essay about a student project to create a shared public space on campus (word types 247, word tokens 446) was generated by ChatGPT as the reading material. Twenty target words (5 verbs; *draw*, *run*, *take*, *hold*, *make*, 5 adjectives; *empty*, *bright*, *sharp*, *great*, *public*, 10 nouns; *idea*, *place*, *change*, *mind*, *line*, *money*, *eye*, *pain*, *difference*, *attention*) were selected from the essay.

### Procedure

The experiment consisted of three parts.

#### *Part I*

All participants performed a lexical relation test (LRT; Test 1). They were instructed to judge the degree of perceived relationship between two target words (100 pairs resulted through the combination of 10 verbs/adjectives and 10 nouns) on a 7-point scale (1: “weakly-related” – 7: “strongly-related”).

#### *Part II*

One week later, participants were divided into three groups according to their scores in Test 1 that were equalized in group means. They read the essay and performed assigned task designed on the basis of the Technique Feature Analysis (TFA) as follows.

Participants in MCQ (multiple-choice question) group answered ten multiple-choice questions about the meaning of the target words (two-word phrase; e.g., *great idea*).

Participants in FIB (fill-in-the blank) group filled in ten blanks in the essay using target words (5 verbs and 5 adjectives) provided with their meanings and usage examples.

Participants in Comp (composition) group completed ten sentences unrelated to the essay with the target words (5 verbs and 5 adjectives) highlighted in the essay.

**Table 1**
*TFA Points Assigned to the Task of Each Group*

| | TFA check items | MCQ | FIB | Comp |
|---|---|---|---|---|
| **Motivation** | | | | |
| 1 | Is there a clear vocabulary learning goal? | 0 | 1 | 1 |
| 2 | Does the activity motivate learning? | 1 | 1 | 0 |
| 3 | Do the learners select the words? | 0 | 0 | 0 |
| **Noticing (Attention)** | | | | |
| 4 | Does the activity focus attention on the target words? | 1 | 1 | 1 |
| 5 | Does the activity raise awareness of new vocabulary learning? | 0 | 1 | 1 |
| 6 | Does the activity involve negotiation? | 0 | 0 | 0 |
| **Retrieval** | | | | |
| 7 | Does the activity involve retrieval of the word? | 1 | 0 | 0 |
| 8 | Is it productive retrieval? | 0 | 0 | 0 |
| 9 | Is it recall? | 0 | 0 | 0 |
| 10 | Are there multiple retrievals of each word? | 0 | 0 | 0 |
| 11 | Is there spacing between retrievals? | 0 | 0 | 0 |
| **Generation (Creative Use)** | | | | |
| 12 | Does the activity involve generative use? | 1 | 1 | 1 |
| 13 | Is it productive? | 0 | 0 | 1 |
| 14 | Is there a marked change that involves the use of other words? | 0 | 0 | 1 |
| **Retention** | | | | |
| 15 | Does the activity ensure successful linking of form and meaning? | 0 | 1 | 1 |
| 16 | Does the activity involve instantiation? | 0 | 0 | 0 |
| 17 | Does the activity involve imaging? | 0 | 0 | 0 |
| 18 | Does the activity avoid interference? | 1 | 1 | 1 |
| **Total Score** | | **5** | **7** | **8** |

TFA points assigned to the tasks were calculated as 5 for MCQ, 7 for FIB, and 8 for Comp according to the theory and presented examples in the TFA literatures (see Table 1). After finishing the task, all participants answered the same LRT (Test 2) as in Part I.

***Part III***

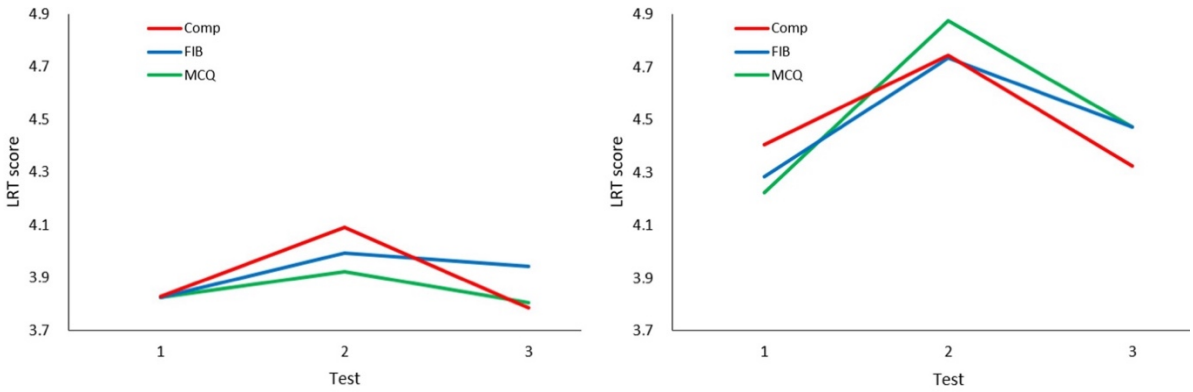
Four weeks after Test 2, all participants answered the same LRT (Test 3) as in Parts I and II.

## Results

### Quantitative Analysis

Ratings for the LRT were averaged for each participant and in each test. Figure 1(a) shows mean LRT score in three tests in each group. Two-way (Group × Test) Analysis of Variance (ANOVA) revealed a significant main effect of the Test factor ($F_{(2,140)} = 4.884$, $p = .009$, $\eta^2 = .065$). Post-hoc analysis showed that the scores were significantly higher in Test 2 than in Test 1 and Test 3. The main effect of the Group factor and the interaction of factors were not significant.

**Figure 1**
*Mean LRT Score in Three Tests in Each Group*




*Note*. (a) For all data (100 pairs) (b) For collocated words data (16 pairs)

Next, an additional analysis was conducted on the data for the 16 out of 100 word-pairs that collocated in the essay. Figure 1(b) shows the results. Two-way (Group × Test) ANOVA revealed a significant main effect of the Test factor ($F_{(2,140)} = 18.629$, $p < .001$, $\eta^2 = .210$). Post-hoc analysis showed that the scores were significantly higher in Test 2 than in Test 1 and Test 3. The main effect of the Group factor and the interaction of factors were not significant.

### Qualitative Analysis (Gephi)

The individual LRT data was analyzed and visualized by Gephi (Bastian et al., 2009), a data-visualization platform, to show how participant's lexical network among target words changed through three tests. Figure 2 shows samples of two participants.

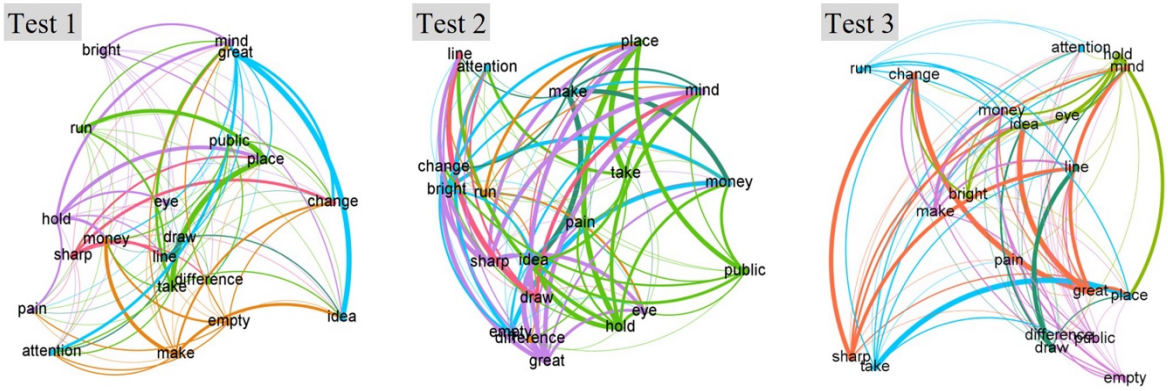
## Discussion

As shown in Figure 1(a), learning effect was clearly demonstrated immediately after the task (Test 2) in all groups. However, the effect did not last for four weeks (Test 3). The same pattern was found in the authors' previous research (Aotani & Takahashi, 2025). Because participants spent only a short time, about 10 minutes in most cases, working on the tasks in these studies, it is likely that long-term learning effects were difficult to achieve.

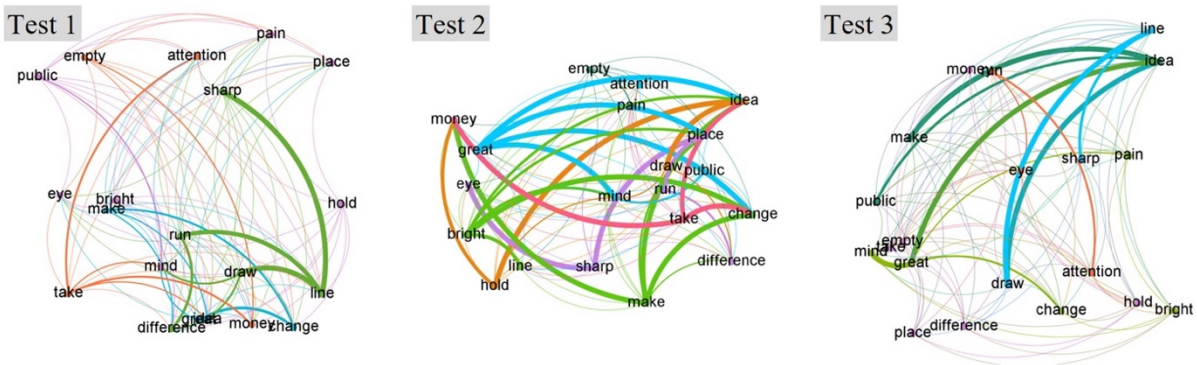
On the other hand, the effects of tasks with different TFA points, which were our main focus of interest, were not observed. According to a TFA point of the task in each group, we predicted that the highest learning effects would be obtained in Comp group, followed by FIB group and MCQ group. Although visual inspection of Figure 1 suggested a trend in this direction, this trend was not statistically significant. There are several possible explanations for this result. First, as mentioned above, the relatively short amount of time spent on the task suggests that participants' engagement with the task may have been insufficient. Second, the differences among the TFA points in three groups (5, 7, and 8 points) may have been too small to result in different learning effects. Finally, TFA is a framework designed for new word learning and may not be directly applicable to facilitating the comprehension of already-known words.

**Figure 2**

*Samples of a Visualized Lexical Network Among Target Words Generated by Gephi*




*Note.* (a) A participant in FIB group (LRT score: Test 1; 1.46, Test 2; 2.21, Test 3; 1.42)




*Note.* (b) A participant in Comp group (LRT score: Test 1; 1.97, Test 2; 2.90, Test 3; 2.00)

The above trend remained unchanged even when the analysis was limited to collocated words in the essay (Figure 1(b)). Again, only the main effect of the Test factor was significant, whereas neither the main effect of the Group factor nor the interaction of factors was significant. Interestingly, a visual inspection of Figures 1(a) and 1(b) suggests that the immediate learning effect appears to be greater, as well as the level of scores itself is higher, for collocated words data than all data. For the participants, the associations between words that they had directly experienced during the task may have been more readily strengthened. However, this finding may also suggest that lexical networks do not readily deepen and expand beyond the directly experienced collocations.

An analysis of the individual data using Gephi yielded some interesting findings. The two samples shown in Figure 2 indicate that the network among the target words became denser from Test 1 to Test 2. In addition, Figure 2(a) suggests that basic verbs such as *make*, *take*, and *hold* developed stronger connections with many nouns in Test 2. Figure 2(b) suggests that the lexical networks of central words in the essay, such as *public* and *place*, became denser in Test 2. However, in both cases, these effects were hardly observed in Test 3. These considerations clarify the importance of investigating individual data as well as group data.

## Conclusion

The present study aimed to identify task characteristics that promote the deepening and expansion of learners' lexical networks for already-known words through extensive reading, using the TFA framework. However, the results of the experiment did not support the

predictions of the TFA, even though the tasks used were conventional ones that have been widely employed in previous research on the retention of newly learned words. Future studies should design new tasks that are tailored to the characteristics of Japanese learners of English and that enable the incidental development of lexical networks through extensive reading. In addition, the TFA framework itself, which was designed basically for the retention of newly learned word, should be reconsidered, and further research should explore the development of a new framework that accounts for the deepening and expansion of lexical networks of already-known words.

## Acknowledgements

This study is supported by Grant-in-Aid for Scientific Research (C) (25K04329) from the Japan Society for the Promotion of Science.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that ChatGPT was used in proofreading and refining the language used in the manuscript, in addition to creating the reading material. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The author further declares that, apart from ChatGPT, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Aitchison, J. (1994). *Words in the mind: An introduction to the mental lexicon* (2nd ed.). Blackwell.

Aotani, N., & Takahashi, S. (2022). Effects of involvement load in extensive reading on lexical relations among already known L2 words. In M. M. Costa, C. M. Almeida, & A. S. Gomes (Eds.), *Education and New Developments 2022* (Vol. I, pp. 519–523). In Science Press. https://doi.org/10.36315/2022v1end116

Aotani, N., & Takahashi, S. (2023). Effects of involvement load of the task on Japanese EFL learners' lexical network changes. In *The Southeast Asian Conference on Education 2023: Official Conference Proceedings* (pp. 635–641). International Academic Forum (IAFOR). https://doi.org/10.22492/issn.2435-5240.2023.51

Aotani, N., & Takahashi, S. (2025). Effects of involvement load on deepening of Japanese EFL learners' lexical network of already-known words. *INTED Proceedings, 1*, 1113–1119. https://doi.org/10.21125/inted.2025.0368

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *Proceedings of the International AAAI Conference on Web and Social Media, 3*(1), 361–362. https://doi.org/10.1609/icwsm.v3i1.13937

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Gohar, M. J., Rahmanian, M., & Soleimani, H. (2018). Technique feature analysis or involvement load hypothesis: Estimating their predictive power in vocabulary learning. *Journal of Psycholinguistic Research, 47*(4), 859–869. https://doi.org/10.1007/s10936-018-9568-5

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*(2), 207–223. https://doi.org/10.64152/10125/66953

Hu, H.-C. M., & Nassaji, H. (2016). Effective vocabulary learning tasks: Involvement load hypothesis versus technique feature analysis. *System, 56*, 28–39. https://doi.org/10.1016/j.system.2015.11.001

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal, 73*(4), 440–464. https://doi.org/10.1111/j.1540-4781.1989.tb05325.x

Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1–26. https://doi.org/10.1093/applin/22.1.1

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press.

Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Heinle Cengage Learning.

Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language, 18*(1), 1–28. https://doi.org/10.64152/10125/66611

Qin, M. (2023). The involvement load hypothesis or the technical feature analysis: Evidence from four EFL vocabulary learning tasks. *International Journal of Education and Humanities, 8*(2). https://doi.org/10.54097/ijeh.v8i2.7501

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language, 15*(2), 130–163. https://doi.org/10.64152/10125/66776

Wilks, C., & Meara, P. (2002). Untangling word webs: Graph theory and the notion of density in second language word association networks. *Second Language Research, 18*(4), 303–324. https://doi.org/10.1191/0267658302sr203oa

**Contact email:** aotani@tokaigakuen-u.ac.jp

# Strengthening Self-Regulation and Executive Functions Through a Value-Based Challenge-Board Framework: Educational Implications for Contemporary Schools

Stephan Drescher, Střední škola UNIQ s.r.o., Czech Republic

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

This paper introduces the Challenge Board Method, a structured pedagogical framework designed to systematically foster self-leadership, responsibility, resilience and core social competences in adolescents. Grounded in research on social learning, executive functions, self-regulation, experiential learning and resilience, the method provides students with a four-year developmental trajectory through incrementally challenging tasks (Explorer, Pathfinder, Pioneer, Leader). The Challenge Board is embedded in a broader personal development programme, which combines a Value Compass (eight core values), practical challenges, Live Labs (theoretical inputs), individual mentoring, portfolio work, expeditions and Vision Quests. Together, these elements form a coherent educational model that supports identity development ("Who am I?" – "Who do I want to become?"), self-leadership and civic responsibility. The method integrates peer mentoring, coaching conversations, portfolio documentation and real-life projects into a unified school-wide system. It requires minimal material resources and is adaptable to various educational settings. The Challenge Board aligns with competence-based curriculum frameworks (e.g. EU Key Competences) and responds to educational challenges associated with motivation, emotional regulation, self-efficacy and social competence. This paper outlines the theoretical foundations, developmental architecture, core structure, implementation processes and observed developmental benefits, and discusses the method's potential contribution to contemporary school development.

*Keywords:* Challenge Board Method, self-leadership development, experiential learning, adolescent resilience, competence-based education, personal development framework

iafor
The International Academic Forum
www.iafor.org

## Introduction

According to OECD (2025), adolescents today grow up in an environment characterised by high cognitive load, fragmented attention, digital overstimulation and limited opportunities for embodied and authentic social learning. Schools report decreasing intrinsic motivation, reduced frustration tolerance and growing difficulties in emotional regulation and social functioning (Gnambs et al., 2023; OECD 2025; Raufelder, D. et al., 2021; Scherrer & Preckel, 2019). These trends underline the need for school-based approaches that deliberately strengthen self-regulation, executive functioning, cooperation skills and personal responsibility, as highlighted in the OECD Learning Compass 2030 and the EU Key Competences for Lifelong Learning.

Universal school-based self-regulation programmes have emerged as a central response to the increasing challenges schools face regarding students' emotional regulation, behavioural self-management and sustained learning engagement. The meta-analysis by Pandey et al. (2018) provides a foundational overview of this field, showing that universal interventions delivered to whole classes or cohorts can produce *small to moderate improvements* in self-regulation, academic performance and behaviour. These findings reinforce the premise that self-regulation is not a static individual trait but a malleable developmental capacity that can be strengthened through intentional pedagogical design and repeated practice within school routines.

Further empirical support comes from more targeted intervention studies. Paananen et al. (2022) demonstrate that structured group-based programmes can enhance attention and executive functions—abilities closely tied to self-regulation—and emphasise their predictive value for both academic achievement and social functioning. Similarly, Lertladaluck et al. (2021) show that an eight-week, school-based mindfulness programme for preschoolers at risk leads to significant improvements in executive functions (working memory, inhibition) and observable self-regulatory behaviour in everyday classroom contexts. Taken together, these studies position executive functions as a key developmental mechanism through which self-regulation programmes exert their effects, even when implemented over relatively short timeframes and within ordinary school routines.

At the level of practice guidance and policy translation, the Education Endowment Foundation (EEF, 2021–2024) synthesises this evidence into accessible frameworks for schools. The EEF highlights self-regulation and executive functions as *high-leverage domains* influencing learning engagement, relationship quality and emotional wellbeing, and stresses elements such as explicit strategy instruction, modelling, scaffolded practice, and metacognitive reflection.

Despite this encouraging evidence base, a critical examination is warranted. First, while effect sizes across studies are significant, they remain modest, and substantial heterogeneity exists regarding programme design, duration and implementation quality. This raises unresolved questions about the *active ingredients* of effective self-regulation programmes and the contextual factors that moderate their impact. Second, many interventions are tested under controlled research conditions with trained facilitators—conditions not easily replicated in everyday school environments characterised by curricular pressure, variable teacher expertise and inconsistent institutional support. Third, there is a broader conceptual risk: when self-regulation is framed predominantly as an individual skill deficit, structural contributors to students' dysregulation—such as overstimulating digital environments, fragmented school schedules, or assessment systems that heighten anxiety—may be overlooked.

Taken together, the literature suggests that universal self-regulation interventions are most promising when embedded within a *holistic school development strategy* that aligns pedagogical practices, relational climates, participatory structures and curricular goals. Rather than functioning as isolated add-ons, such programmes should be integrated into a coherent educational ecology that supports students in developing executive functioning, emotional regulation, cooperative capacity and personal responsibility—competences increasingly recognised as essential for navigating contemporary learning environments.

The OECD Learning Compass 2030 articulates a long-term educational vision in which students develop the capacity to navigate complexity with agency, responsibility and resilience. Central to this vision is the idea that learners must be equipped not only with academic knowledge, but with self-regulation, social competence, ethical awareness and the ability to co-create solutions in uncertain environments (OECD, 2019–2024). The accompanying Student Agency for 2030 Concept Note further emphasises that agency is inherently relational and value-based: it requires the ability to set meaningful goals, regulate one's behaviour, take responsibility for actions and contribute constructively to the community (OECD, 2019). These priorities align closely with the EU Key Competences for Lifelong Learning (Council of the European Union, 2018), which frame personal, social and learning-to-learn competences—along with citizenship competence—as foundational capacities enabling individuals to act autonomously, collaboratively and responsibly across the lifespan.

Within this broader policy landscape, a systematic pedagogical framework such as the Challenge Board can be understood as a practical, school-level response to these international aspirations. By combining structured self-regulation tasks, cooperative challenges, value-based reflection and progressive responsibility-taking, the method operationalises many of the competencies highlighted by OECD and EU frameworks. Moreover, its multi-year developmental architecture explicitly cultivates student agency by engaging learners in goal-setting, monitoring, mentoring dialogues and authentic experiences that extend beyond the classroom. In this sense, the Challenge Board does not merely align with global competency frameworks—it provides a concrete methodological pathway through which schools can translate the high-level ambitions of the Learning Compass into sustainable everyday practice.

Many existing approaches—such as Social and Emotional Learning (SEL) programmes, project-based learning or outdoor education—provide valuable experiences but are often implemented as isolated projects or short-term interventions. Schools frequently lack a systemic, multi-year and scalable instrument that:

- embeds challenge-based learning into the regular school routine,
- links theory and practice,
- makes personal development visible and documentable,
- and remains manageable for teachers.

The Challenge Board Method was designed to address this gap. Personal development is conceived as a four-year, systematically structured process that guides students from self-knowledge towards self-leadership and responsibility:

- **Phase 1 – Discovering Identity (Years 1–2)**
  *Who am I? What are my strengths? How do I function as a person?*
  → Focus: self-awareness, strengths analysis, basic competences.
- **Phase 2 – Shaping Identity (Years 3–4)**
  *Who do I want to become? How do I want to take responsibility?*
  → Focus: life vision, values, responsibility, self-leadership.

This developmental logic is operationalised through six tightly interlinked elements:

1. the UNIQ Value Compass (eight core values),
2. practical challenges (Challenge Board),
3. theoretical and reflective inputs (Live Labs),
4. individual mentoring and coaching,
5. portfolio work as a "development book,"
6. experiential learning spaces (expeditions, Vision Quests, projects, career guidance).

Within this ecosystem, the Challenge Board serves as the central instrument that systematically brings students out of their comfort zone in a safe, structured and pedagogically guided way, and provides a clear scaffold for personal growth.

## Theoretical Framework

### Social Learning and Peer Modelling

The Challenge Board draws on Bandura's Social Learning Theory, particularly the role of modelling, vicarious reinforcement and self-efficacy. Older students act as mentors and tutors, enabling younger peers to observe strategies for perseverance, emotional regulation and problem solving in authentic contexts. Research on peer-assisted learning shows that such near-peer models can significantly enhance motivation, skill acquisition and social belonging.

### Executive Functions and Self-Regulation

Research in developmental psychology and educational neuroscience shows that executive functions—planning, inhibition, working memory, sustained attention—are strong predictors of academic and life outcomes. The Challenge Board method trains these capacities through:

- structured goal setting,
- micro-challenges,
- mid- and long-term projects,
- regular coaching conversations,
- and systematic portfolio reflection.

Rather than teaching self-regulation as an isolated content theme, the method embeds it into **real, meaningful tasks**.

### Resilience, "Out of Comfort Zone" and Stress Inoculation

Resilience is understood as a dynamic process supported by protective factors such as supportive relationships, positive expectations and repeated experiences of overcoming manageable difficulties. The Challenge Board provides a sequenced exposure to challenge: from short micro-challenges to multi-day expeditions and, ultimately, optional long-term expeditions in unfamiliar environments.

Within this design, students intentionally leave their comfort zone and enter a "stretch zone"—situations that are demanding but not overwhelming. Progressive challenge, combined with mentoring, peer support and structured reflection, operationalises principles of stress inoculation: controlled exposure to difficulty builds confidence, frustration tolerance and coping competence.

### Experiential Learning and Identity Formation

Kolb's experiential learning cycle—experience, reflection, conceptualisation, planning—is systematically implemented through:

- concrete challenges in four domains,
- reflection in Live Labs and coaching,
- conceptual inputs on brain function, emotions, communication and ethics,
- and planning of next steps in mentoring sessions.

Across four years, this cycle supports identity work: *"Who am I?"* (Years 1–2) and *"Who do I want to become and what life do I want to lead?"* (Years 3–4). Vision Quests and expeditions provide intensified experiences outside everyday school life that deepen this identity work.

### Competence-Oriented Education and Curriculum Alignment

The UNIQ personal development programme is aligned with the Czech national curriculum (RVP) and the school-based curriculum (ŠVP). The Challenge Board contributes directly to key competences such as learning to learn, problem solving, communication, social and personal competence, civic competence and work-related competence. These links are made explicit in the school's ŠVP by connecting:

- **Value Compass** → school ethos and value education,
- **Challenge Board** → project-based, competence-oriented learning,
- **Portfolio** → learning-to-learn competence and metacognition,
- **Live Labs** → explicit instruction in social, emotional and civic competences,
- **Expeditions and Volunteering** → civic and environmental responsibility.

### Developmental Logic of Personal Growth (Phase Model)

The four-year programme follows a coherent developmental logic:

- **Phase 1 (Years 1–2): Discovering Identity**
  Live Labs focus on self-knowledge (brain function, temperament, strengths and "shadow sides," mindset, emotional regulation). Challenges are smaller in scope; expeditions offer 1–2 nights away from home. Students learn basic self-observation, reflection and first steps in responsibility.
- **Phase 2 (Years 3–4): Shaping Identity and Future**
  Live Labs shift towards communication, relationships, society, ethics and finally future and self-reliance (financial literacy, entrepreneurship, life vision, time management). Challenges increase in complexity and social responsibility. Expedition formats include multi-day and optional long-term expeditions, in which students plan and implement projects with high autonomy.

The Challenge Board is the structural backbone that connects these phases and translates abstract development goals into concrete, assessable activities.

## Method Description: The Challenge Board

### Structural Overview

The Challenge Board is a four-year developmental system with four levels:

1. **Explorer (Year 1)** – Orientation and first steps outside the comfort zone

2. **Pathfinder (Year 2)** – Focus, perseverance and consolidation
3. **Pioneer (Year 3)** – Self-organisation and self-efficacy
4. **Leader (Year 4)** – Responsibility for others and leadership

Each level involves:

- **Micro-challenges** (seconds to one hour)
- **Mid-term challenges** (1 day – 1 week)
- **Long-term challenges** (weeks to one school year)
- **Four major annual challenges** in the domains:
    - Volunteering
    - Physical Challenge
    - Skill Challenge
    - Expedition

Participation is formally voluntary but strongly recommended. Experience has shown that intrinsic motivation and ownership are crucial for deep learning; therefore, students are invited—but not forced—to engage. Students joining the school later (e.g. in Year 3) start with the Explorer level and can progress at their own pace; access to the Leader level is possible with mentor and peer-group approval in cases of exceptional engagement.

**The UNIQ Value Compass as Integrative Framework**

At the heart of the Challenge Board lies the UNIQ Value Compass, consisting of eight core values:

1. Courage
2. Focus
3. Patience
4. Cooperation
5. Creativity
6. Resilience
7. Aesthetics
8. Responsibility

These values serve as long-term developmental dimensions rather than abstract ideals. In practice:

- Every challenge is assigned to at least one value.
- Portfolio reflections are structured along these values.
- Mentoring conversations use the Value Compass to talk about strengths, blind spots and developmental edges.
- Live Labs pick up the values as thematic anchors (e.g. “Courage and decision-making,” “Patience and process thinking,” “Responsibility and ethics”).

This design ensures that personal development is not driven by grades but by inner growth and responsible action. Students progressively become aware of their personal values, recognise strengths and development areas and take increasing responsibility for their own trajectory.

## Types and Domains of Challenges

Within each level, the Challenge Board differentiates several challenge types:

- **Micro-challenges**: short tasks embedded in the school day (e.g. short physical exercises, mini-presentations, small "courage experiments").
- **Mid-term challenges**: activities lasting from one day to one week (e.g. new routines, mini-projects).
- **Long-term challenges**: extended undertakings over several weeks or months (e.g. volunteering projects, larger sports goals, creative or technical projects).

All challenges can be student-initiated (supporting autonomy and self-determination) or teacher-initiated (offering developmental impulses and scaffolding). Major challenges in the domains Volunteering, Physical, Skill and Expedition are planned and monitored at the beginning of the school year and revisited in mentoring sessions.

Extensive catalogues of possible challenges (e.g. in physical, skill or volunteering domains) are compiled in manuals for teachers and students and serve as inspiration, not as prescriptive lists.

## Role Distribution: Teachers, Mentors and Student Tutors

To keep the system manageable and to cultivate a culture of shared responsibility, roles are clearly defined:

- **Teachers / Mentors**
  - focus on the four major challenge domains,
  - conduct mentoring and coaching conversations,
  - ensure safety in physical and expedition contexts,
  - approve annual goals and monitor progress,
  - oversee portfolio quality and award decisions.
- **Student Tutors**
  - support peers with micro-challenges,
  - confirm completed micro-challenges in the portfolio,
  - provide low-threshold, peer-level encouragement and feedback.

This peer mentoring system strengthens social cohesion, builds leadership competence among older students and significantly reduces administrative burden for teachers.

## Coaching and Live Labs

Two core formats anchor reflection and theory:

- **Monthly coaching conversations**
  Each student has at least one individual coaching session per month, typically embedded in a fixed "personal development block." These follow a standard structure:
  1. Review: What have I done?
  2. Reflection: What have I learned?
  3. Evaluation: What went well or poorly, and why?
  4. Planning: What are the next steps?
  5. Portfolio: Update and confirmation.

- **Live Labs (2× per month in Year 1, then at regular intervals)**
  Live Labs provide theoretical and reflective input on themes such as brain function, personality, strengths and shadows, emotion regulation, communication, conflict resolution, digital media, civic engagement, financial literacy and life vision. They connect academic content with the Value Compass and the Challenge Board, ensuring that experience and concept are tightly linked.

**Portfolio as "Development Book"**

Each student maintains a personal development portfolio, which documents:
- challenge planning,
- evidence (photos, videos, artefacts),
- written reflections and Talking-Head videos,
- mentor feedback,
- annual summaries and award status.

The portfolio functions as a longitudinal record of personal development and can, if desired, be presented in later application processes to provide a nuanced picture of competences, engagement and learning attitude. Teachers have clear but limited documentation responsibilities: they confirm major challenges, key reflections and award decisions in both the portfolio and the student record.

**Award System and Assessment Logic**

Assessment follows a deliberately non-graded, binary system:
- ✔ *Passed*
- ✖ *Not passed*

For each school year, students may receive an award corresponding to their level (Explorer, Pathfinder, Pioneer, Leader). An award is granted when:
- the major challenges have been started and completed,
- the portfolio is complete and reflective,
- coaching sessions have taken place regularly,
- the student has demonstrably taken responsibility for their own development,
- minimum standards agreed in the annual start-of-year conversation are met.

This design emphasises growth and responsibility rather than external performance comparison and is fully compatible with competence-oriented assessment standards.

**The Four Major Challenge Domains**

***Volunteering Challenge***

Volunteering challenges foster empathy, social responsibility and civic engagement. Complexity and scope increase across levels—from single days of engagement to assuming long-term responsibilities or leading projects and younger student groups. A structured concept of "responsibility areas" (e.g. school garden care, library, school café, local community support) anchors volunteering in everyday school life and the local community.

### *Physical Challenge*

Physical challenges aim at body awareness, endurance, discipline and mental resilience. They range from micro-challenges (e.g. short stabilisation exercises or coordination tasks) to long-term sports goals and participation in competitions. Advanced students may take on tutor roles, leading younger peers in physical micro-challenges.

### *Skill Challenge*

Skill challenges promote talent development, learning culture and self-efficacy in creative, musical, linguistic, technical, practical or social domains. They may include, for example, learning a musical instrument, developing a technical or craft-based skill, creating an artistic work, acquiring a new language competence or organising a practical project. Over four years, students progress from initial exploration to visible products or performances and finally to tutor roles in which they teach younger peers and assume responsibility for planning and reflection. In this way, individual interests and talents gradually develop into demonstrable competences that can be shared with and contribute to the school community.

### *Expedition and Long-Term Expedition Challenge*

Expeditions are deliberately designed "out of comfort zone" experiences in nature or urban environments. They increase over the years from teacher-led, short trips with one overnight stay to largely self-organised multi-day expeditions and, optionally, long-term expeditions (1–2 weeks, ideally abroad) with high levels of planning responsibility (budget, travel, safety concept, learning tasks, documentation).

Vision Quests are integrated into expeditions and provide focused spaces for identity and future reflection (e.g. "Who am I?", "What are my strengths and weaknesses?", "Who do I want to become?"). Safety standards, roles and procedures are documented in an Expedition Guide and follow clear planning, risk assessment and aftercare protocols.

## Educational Impact

### Self-Leadership and Responsibility

The Challenge Board systematically trains autonomous decision-making, initiative and responsibility. Students plan, implement and reflect on challenges over multiple years, learning to own their decisions and consequences before they are asked to take responsibility for others (Leader level).

### Executive Functions and Self-Regulation

The repeated cycle of planning, acting, monitoring and adjusting strengthens executive functions such as planning, inhibitory control and working memory. The combination of micro-, mid- and long-term challenges allows students to practise self-regulation in different time horizons.

### Social Competence and Peer Collaboration

Structured peer mentoring, responsibility areas and cooperative challenges foster communication, empathy, conflict resolution and teamwork. Students experience themselves as part of a community in which everyone's contribution matters.

### Resilience, Comfort Zone and Well-being

By intentionally stepping out of their comfort zone in manageable steps (physical challenges, social exposure, expeditions, Vision Quests), students build frustration tolerance, coping strategies and resilience. Reflection in mentoring and Live Labs supports emotional processing and contributes to psychological well-being.

### Motivation, Identity and Future Orientation

The level structure (Explorer → Leader), the portfolio as a visible growth narrative and the integration of career guidance and Vision Quests strengthen intrinsic motivation and future orientation. Students do not only ask *"What do I want to study?"* but also *"What kind of person do I want to become and how do I want to contribute?"*

## Practical Implementation

### Time Structures

The method is anchored in the school timetable through:

- a monthly 2-hour personal development block for Challenge Board sessions, coaching and portfolio work,
- regular Live Labs (e.g. two per month in Year 1),
- expedition windows (e.g. winter and spring expeditions in Year 1),
- weekly open work phases for individual project work and micro-challenges.

### Standardised Processes

To ensure consistency and reduce organisational load, processes are standardised:

- registration of Challenge Levels (Explorer–Leader) in autumn,
- annual planning conversations and mentor approval,
- monthly coaching sessions and portfolio updates,
- clear procedures for micro-challenge registration and confirmation by student tutors,
- structured processes for Live Labs and expeditions (goal, method, reflection, documentation),
- annual award decision and documentation in student records.

These processes are summarised in the UNIQ Teacher Handbook and provide a transparent framework for teachers and students.

### Alignment With School Governance and RVP/ŠVP

The Challenge Board system is formally embedded in the Czech school's ŠVP as the central instrument of the *personal development module*. It is explicitly linked to:

- school values and ethos (Value Compass),

- competence- and project-based learning,
- assessment principles (pass / not pass),
- individual support and inclusive education.

## Evaluation and Observations

Initial observations in pilot implementations indicate:
- increased self-confidence and autonomy,
- improved emotion regulation and resilience,
- strengthened peer cooperation and sense of belonging,
- higher engagement in school tasks,
- improved reflective competence and metacognition,
- more constructive and meaningful teacher–student coaching conversations.

Portfolios and student feedback suggest that the combination of Value Compass, Challenge Board, Live Labs and expeditions is experienced as coherent and meaningful. Teachers report that behavioural issues decrease while students' initiative and sense of responsibility rise. For external stakeholders (e.g. employers, universities), the portfolio provides a nuanced documentation of competences beyond grades.

Systematic empirical studies (e.g. pre–post designs, comparison groups) are planned to quantify these effects and to analyse long-term developmental trajectories.

## Discussion

The Challenge Board shares features with existing frameworks such as the Duke of Edinburgh Award, SEL programmes and outdoor education models, but is distinctive in several ways:
- its integration into a four-year, whole-school personal development programme,
- the central role of the Value Compass as a cross-cutting reference for all elements,
- the multi-layered challenge architecture (micro, mid, long-term; four domains),
- the systematic use of peer mentors and tutors,
- the strong embedding in curriculum and assessment frameworks (RVP/ŠVP).

Challenges include ensuring consistent mentor training, calibrating expectations across classes and maintaining a balance between standardisation and individual flexibility. Nonetheless, the method offers a robust, scalable and low-threshold pathway for schools seeking to systematically cultivate self-leadership, resilience and social responsibility.

## Conclusion

The Challenge Board Method represents a comprehensive, developmentally sequenced and institutionally embedded framework for personal development in secondary education. By combining challenge-based learning, value education, reflection, peer mentoring and competence-oriented assessment, it enables adolescents to grow beyond their comfort zones and to become stable, reflective and responsible personalities.

The method demonstrates how schools can:
- align personal development with national curriculum frameworks,
- integrate theory and practice,
- and structure "out of comfort zone" experiences in a safe, pedagogically guided way.

Further empirical research is recommended to evaluate developmental outcomes, refine implementation standards and explore transferability to other school contexts.

## Acknowledgements

I would like to express my sincere gratitude to the Karel Janeček Foundation for supporting the scientific development and research of this method. My heartfelt thanks also go to my four children and to all children who participated in the pilot phase of the project. Their continuous engagement with the method has provided invaluable joy and insight, and has deeply inspired me to refine the approach and make it accessible to a wider community of learners.

## References

Council of the European Union. (2018). *Council Recommendation of 22 May 2018 on key competences for lifelong learning* (2018/C 189/01). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.C_.2018.189.01.0001.01.ENG

Education Endowment Foundation. (2021–2024). *Guidance reports on executive function and self-regulation.* EEF / EdResearch.

Gnambs, T., & Hanfstingl, B. (2023). In Alivernini, F., Manganelli, S., & Mallia, L. (Eds.), *Understanding and supporting the motivation of students.* OECD Publishing.

Lertladaluck, K., Suppalarkbunlue, W., Moriguchi, Y., & Chutabhakdikul, N. (2021). School-based mindfulness intervention improves executive functions and self-regulation in preschoolers at risk. *The Journal of Behavioral Science, 16*(2), 58–72. https://so06.tci-thaijo.org/index.php/IJBS/article/view/249238

OECD. (2019). *Student Agency for 2030: Concept Note.* OECD Publishing.

OECD. (2019–2024). *OECD Learning Compass 2030: Concept Notes & The Future of Education and Skills 2030.* OECD Publishing.

OECD. (2025). *How's Life for Children in the Digital Age?* OECD Publishing.

Paananen, M., Aro, T., Korhonen, J., & Tapola, A. (2022). School-based group intervention to improve attention and executive functions in children. *Frontiers in Psychology*, *13*, 824523. https://doi.org/10.3389/fpsyg.2022.975856

Pandey, A., Hale, D., Das, S., Goddings, A.-L., Blakemore, S.-J., & Viner, R. M. (2018). Effectiveness of universal self-regulation–based interventions in children and adolescents: A systematic review and meta-analysis. *JAMA Pediatrics*, *172*(6), 566–575. https://doi.org/10.1001/jamapediatrics.2018.0232

Raufelder, D., Jagenow, D., Hoferichter, F., & Drury, K. M. (2021). The role of the learning environment in adolescents' academic motivation. *Learning and Individual Differences*, *90*, 102046. https://doi.org/10.1007/s11031-021-09879-1

Scherrer, V., & Preckel, F. (2019). The decline of academic motivation during adolescence: An integrative review of longitudinal studies. *Educational Psychologist*, *54*(4), 250–270. https://doi.org/10.1080/01443410.2015.1113236

# The Price of Progress? Assessing the Impact of Artificial Intelligence on Systems Engineering Education

Kieran Holling, Cranfield University & Defence Academy, United Kingdom
Steve Barker, Cranfield University & Defence Academy, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

The advent of Artificial Intelligence (AI), especially Generative AI (GAI), could have both advantages and disadvantages for the education sector. In a positive sense, GAI allows faster more comprehensive searching for information and may facilitate the development of models and concepts at a higher tempo. In a negative sense, however, AI generated content, which might pass as the work of a human, could be submitted by unscrupulous individuals as their own work. This threat to original insight and thought is a challenge to the integrity of educational qualifications. Beyond content generation, however, GAI could pose more existential questions for the andragogical process and provision of education. In the field of Systems Engineering (SE) education has, to date, focused on the application of concepts, principles, modelling techniques and related ideas to finding the best possible solution to problem situations in engineering and business. The advent of GAI could have the effect of changing this paradigm; does such technology confer SE ability on those with the ability only to construct the cleverest "prompts"? Can GAI under the right circumstances be used to determine the most appropriate models and approach to solution development and, if so, will higher-order functions be lifted from the human "operator"? If so, where does the human add value in the subject? In the light of these questions, will GAI fundamentally affect how SE should be taught? This paper assesses these questions and proposes andragogical ideas for the constructive use of GAI techniques in SE education.

*Keywords:* artificial intelligence, systems engineering, education

iafor
The International Academic Forum
www.iafor.org

## Introduction

The impact of Artificial Intelligence (AI) is a subject which is of interest or concern to all of us in education to some degree, whether it be positive anticipation of the teaching benefits that AI might offer or it is concern for the potential impact on our students' cognitive abilities.

In our own subject of Systems Engineering (SE), we consider that AI will be used in aspects of practice and of education, but we consider that as SE is so dependent on novel ideas and innovative solutions, an overreliance on AI might result in limited creativity—it might result in new arrangements of existing elements to update an existing solution rather than finding a radical new solution or even identifying a new root cause of the original problem.

To understand why that might be, it is necessary to understand a bit more about our subject and Systems Engineering is not a name that does us any favours. In fact, it may well put many people off. Certainly, the subject deals with *systems* but *engineering* in our case is in the broadest sense of the word—"purposefully designing or arranging" might well describe what we do and that can be applied to all manner of systems from huge governmental systems down to hard product-systems. We are in the business of problem solving; of "engineering" some form of solution to satisfy a user's need.

Importantly, SE is a pan-disciplinary approach; Systems Engineers seek to bring together teams with a range of skills to solve the problem at hand. As SE rises from a number of disciplines, it has inherited its characteristics from a number of forebears. Some of that heritage is use of graphical methods to articulate ideas about the system we are investigating. Again, this is a creative activity and the impacts of that we shall consider later.

Overall, SE is an explorative and creative activity with a large human component, and these factors might well be key to how AI is able to be used to best effect.

In addressing the potential implications of AI for SE education, this paper will: identify why the rise of AI might present a problem for education in general and Higher Education (HE) in particular; look at the particular nature of SE education and consider whether there are specific issues as a result; consider the nature of masters level education and the challenges posed by AI for higher order thinking at this level; conclude by considering the specific threats and opportunities raised by AI in SE education. A precis of this paper, in this same structure was delivered as a recorded presentation in the European Conference on Education in London 9–13 July 2026.

### Genesis of the Problem

Whilst at the start our study, we concerned ourselves with the challenges facing SE education, even the most cursory exploration of the problem space shows that AI is not only an issue for SE nor only for the education sector—it is something which is impacting every aspect of life in the developed world.

There is an idea in systems theory that a change in one part of a system causes change in all other parts of the system (Hall, 1956, as cited in Chestnut, 1967). This is an important idea as we also need to recognise that most systems are "open"—that is to say that they interact with their environments. We can easily see that this means that a change in business practice or students' "schema" of prior experience will have an impact on education and the "customer

need" that we need to fulfil. To understand the needs of our students, then, we need to consider both the motivations of the students in entering HE and also the expectations at entry, based on their prior interaction with the educational and wider societal *(system) environment*.

There have been many societal changes in recent years and there has, in the UK, been a huge increase in the proportion of people going on to Higher Education. Uptake of university places rose in the UK by 174% between 2002 and 2002 (Carasso & Plume, 2023). This is mirrored around the world and "UNESCO's first Higher Education Global Trends Report shows that the number of students enrolled in higher education worldwide has more than doubled over the past two decades, reaching 269 million in 2024" (UNESCO, 2026). In the UK at least, this increased level of education at input to the workforce has created intense competition for jobs. As a result, students are looking for a direct improvement in employability as an outcome of studying for a degree.

Students come to HE already well experienced in the use of AI; they are what has been termed "digital natives". Staff in HE are, by comparison, more likely to be "digital immigrants". They are, by comparison, likely to be coming to AI and other new ways of working long after they learned their craft (Prensky, 2001). Staff are more likely to view education as valuable for its own sake, rather than as a means to get a job (Naylor et al., 2021) and in this context may view the use of AI as a threat to deeper learning rather than an enabler to acquisition of a workplace advantage. So, the prevailing debate on the use of AI in education might be: Have the changing needs of students and universities created a mismatch in expectations?

So, we have established a reasonable proposition that in a world which moves at pace and in which many consider that AI will always be available in the workplace, students are quite naturally focussing on outputs, not journeys. This is entirely understandable as almost every other aspect of daily life encourages students to "outsource" activities to AI assistance—from visually searching an item using the mobile phone camera to asking for rapid summaries of all manner of information which would previously have taken hours or days to research.

So, with the incredible capability that AI provides, and noting that there has been a widespread concern about its use in HE, are academics playing "King Cnut"—the ancient British king who tried to hold back to tide by force of his will (Henry of Huntingdon, 1996, orig. pub. 12$^{th}$ Century) and is forever cited as an illustration of futility. Are they setting themselves up as "Gatekeepers" when all can see that the tide of modernisation is about to sweep them away?

To put this in context and to avoid dying on a hill of their own making, academics need to consider what is important in education. Most recent concern in academia seems to have focussed on the ability of students to produce acceptable responses to assignments by using AI and, as a result failing to learn the material that those assignments were intended to assure. This neatly exposes the challenge faced here—teachers in HE are worrying about the continued ability of current assessment methods to *substantiate* learning as a result of a change in our *system* which has already occurred, rather than worrying about the bigger issues of getting students to learn or even whether the intended learning is still appropriate for students in the changed system.

At its worst, we could see this as a change in demand from our students; a desire to be trained for the workplace rather than being educated for a changeable future. But, if the system has been changed and pre-employment training is in fact what the intended beneficiaries of our service require, should we acquiesce? Alternatively, ought we to revisit education beyond

assessment and focus on teaching and learning and consider whether AI can assist there, and whether they will remain broadly applicable in a way that will see education in an unchanged set of disciplines or whether education might be more appropriate in some subjects or to a changed target audience.

Finally in considering the challenge posed to education in general (before considering SE specifically) it is worth considering the danger of adopting a parochial, even defensive viewpoint. There is of course, a temptation to see one's own subject as being special; uniquely complicated and exquisite. At this point we might do well to heed the admonishment in *How Big Things Get Done* by Flyvbjerg and Gardner (2023), in which they point out that "your project is not unique" and that if one looks hard enough, it will fit into a "class" of project from which lessons can be drawn. Bearing this in mind, it is probably true that there are points which can be read across between SE education and that in other subjects, if we can see common factors.

**Systems Engineering Education**

In understanding the SE educational offering across Europe and the US it is interesting to note that there are many domain-centric systems engineering courses and relatively few system-centric SE courses. By that I mean that there are lots of courses or courses which deal with systems of some specific type, for example Computer Systems Engineering rather than "pure" SE courses based on systems theory and specific techniques like Model Based Systems Engineering. As previously discussed, SE as a problem-solving discipline as opposed to traditional engineering.

Equally there are, in Europe at least, few bachelors' degrees in pure SE and more MSc programmes. Where there are BScs in Systems Engineering in the UK, they are in essence broad based "hard" engineering courses with a Systems "overlay", though with a 3-year full time or 4 year "Sandwich" course, there is plenty of time to cover SE content which has, 'til now been delivered as an MSc "top-up" to current practitioners. For the wider adoption of SE, then, this looks like a good thing, and those courses will face the same problems of protecting original thought and critical thinking in the same way that the "systems centric" courses do.

Experience in Cranfield University—which runs one of the most established SE courses in the UK—many, if not most students come with workplace experience of SE. Those students also report that many of their contemporaries have no formal education in SE but have come across from allied disciplines (in line with our original "pan-disciplinary" definition). So, anecdotally, a significant proportion of SE practitioners have no formal *education* in SE, though may have had *training*, especially in the use of specific software tools.

Both in Europe and in the US, a surprising proportion of establishments delivering SE are linked to government and defence. This is, perhaps, unsurprising as SE theory developed largely in the aerospace and defence sectors. We do not perceive that this will impact how AI in SE education needs to be handled, but there are likely to be other educational anomalies of courses in Military Universities which would bear investigation, including easier control in start standards in the student body, for instance, though that is out of scope for our study.

Reading course descriptions from universities on websites shows many courses both in Europe and the US stress the practical aspects of their courses with "applied", "project", "group work" and "hands on" being common phrases. Another term which appears regularly is "innovative".

This is hardly surprising as the problem-solving nature of the discipline and in the common customer groups of aerospace, defence to some degree transport systems are all seeking to exploit cutting-edge technology in their engineered systems.

Importantly, SE involves a significant degree of analysis even to the point where the Systems Engineer defines for themselves what elements constitute the system—boundaries are not always defined in the real world and ruling elements in or out fundamentally affects the questions asked and the solutions offered. This is *creative* activity.

Finally, Systems Engineers often communicate their ideas visually—the syntax of the languages is clear, but the product, like writing, is a thinking process where ideas and diagrams are developed simultaneously. Again, this creation, often collaborative, is a cognitive process. The experience on Cranfield courses to date is that students have experimented with the use AI to generate SysML diagrams. But have had little success. The diagrams produced contain the elements specified in prompts, but the illustration of how the elements interact does not work, even when specified line-by-line. Of course this may improve over time, but the point previously made about writing, that thinking the text into being is a creative activity in itself, might equally be read across to drawing the SysML diagrams—getting a visually appealing result that we do not understand will still not get us closer to the truth.

Taking these factors together, we can see a building pattern of relationships with a dynamic human dimension, and a need to encourage and *practice creative thought*.

**Artificial Intelligence and Masters' Level Education**

There is no doubt that AI can answer questions quickly and *convincingly*. Those who possess a talent for "Prompt Engineering" can research topics quickly and effectively. This is now part of everyday life for most people involved in the field of HE. However, what has caused concern in many instances, is that convincing is not necessarily the same as correct and this raises the issue of trustworthiness. Applied to an extensive canon of knowledge AI can create valuable and accurate summaries of a subject area, saving a great deal of time. However, where the body of knowledge is less extensive results are less reliable and one known problem is "hallucination" which refers to generating imagined or incorrect content (Alqahtani et al., 2023).

AI tools offer promising formatting, style-checking and editing solutions. Used sparingly abilities can help a student to concentrate on their own research of the topic area and on critical analysis. However, overuse can be disadvantageous as the act of writing is inextricably linked with thinking (Hurley, 2026) and, by implication in an educational context, learning. Where the use of AI is intended by the student to lighten the burden of thinking, learning is likely to suffer and as Alqahtani et al. observed in their 2023 article "(Relying) on AI-generated text to start writing research outside of your expertise may limit your thought process…, potentially hindering your creativity and unique ideas". This same sentiment is echoed in creative industries and, reflecting on the impact on literature, Dr. Clementine Collet noted in her 2025 paper that AI had a place "simplifying non-creative tasks" or that it might be able "to do our admin tasks so that we can focus on the creative aspects of the industry" (Collet, 2025).

The speed of change will be hot and, in many cases has already happened. In the experience of the authors, AI adoption across the student body has been balanced across cohorts, with no particular business sectors seeming to lead or lag. Anecdotally, businesses which resisted AI

adoption in 2022/3 to prevent erosion of intellectual capital gave ground in less than a year and did not founder, as both customers and competitors were changing with them. The ability to use AI effectively is assumed in the student body and the acceptability of use in an educational setting, other than for written assignments, is assumed.

What we do not see changing in SE is the requirement for original thought to both conceive the system to be worked on and to identify novel solutions. This latter point is important as current AI works on trawling previously generated knowledge to apply a similar solution. In this case, finding novel solutions is not a possible use of current AI. This then is where SE finds its "class" (referring back to Flyvbjerg and Gardner's idea in para. 2, p. 4)—use of AI in practice and, therefore, most likely in education will be in common with creative industries (not just artistic ones) which rely on creating completely novel ideas for their ability to progress.

In SE teaching practice at Cranfield University, we have not been immune to concerns regarding use of AI by students. Initially, we observed similarity in approach across class group essay submissions. Whilst an initial assumption might be conferring between student, a simple investigation revealed convergence as a potential indicator of Generative AI use—students using similar prompts (cued by the question as set) will generate very similar answers from the same available data—AI is working on simple probability, not original perspective. This is easily addressed with students by raising the issue that the similarity can be detected. However, this still leaves the question as to whether essay style assignments are tenable as the majority form of summative assessment. Furthermore, there is a potential cost impact of employing AI detection software, not least in the potential costs of dispute resolution.

In a similar vein, there have been instances of referencing which appeared to have facets of "Hallucination" mentioned earlier when addressing the potential use of AI in research. Imagined "References" combine known authors with credible titles, but attempting to follow links can lead to dead-ends or tangentially relevant material. Again, this issue is easily remedied by application to detailed marking and enforcement of the rules. However, it is worth consideration of whether the conflicting requirements of academic development and the students' desire for career enhancing skills are at odds here. In a time where there is no shortage of information and many students might consider that getting to an answer is more important than the journey in getting to it, perhaps the appetite for lengthy research is waning. In that case, perhaps there is a place for skills based short courses for students who prefer that approach, leaving only the more academically inquisitive to undertake long-form degrees. This is counter to the direction of HE in recent years but in a changed context and an era of cost pressure for students, perhaps change is inevitable.

This last point brings us back to the issue we identified when considering "mindset" at the beginning. In SE terms, the environment in which we are operating has changed (advent of ubiquitous AI) and the and our "capability" to achieve our goals (assessment to assure learning) has been degraded. As a result, we have options. Two obvious ones are to: a) reinforce the current assessment method to achieve assurance despite AI or b) to change assessment methods to favour deep understanding and changing our "study skills" teaching to encourage exploitation for the new and powerful AI toolset. It might be hoped that this would support development of a problem-solving mindset—the affective domain of SE learning.

It would be tempting to fall back on old assessment methods which defeat the use of AI; a return to examinations. This is, however, fraught with issues and recent Cranfield experience

is that students are resistant and this to the point of protesting to staff that exams, even supported with access to the students own physical (hard copy) notes were unrealistic in the modern era, where open access to online search will always be available. Add to that the observation that exams are not well suited to the higher orders of Bloom's Taxonomy (Bloom, 1956), and traditional exams may not be the panacea that one might have hoped.

Another option is oral exams or Vivas. Cranfield University currently use these on some courses, though no longer in SE. There are issues with the time taken to conduct significant oral exams, but Vivas to support a written submission are a viable option. Importantly, they address the observation that students who use AI for *writing* often have poor recollection of their argument and are unable to quote their own submissions directly. A viva could, therefore, be a way of detecting use of AI in creating the original, written assignment. Much more than this, however, and what we as educators might more properly be interested in, preparation for a viva might encourage the deeper learning we are trying to foster. However, such a change is significant and comes at cost—both financially and on a human level and such a change would need proper consideration.

What we have concentrated on thus far has largely been teaching and assessment. If we are able to design assessment so that it focusses on understanding and an ability employ ideas in creative ways in real time, then we could incentivise students to use AI to learn rather than answer. As far back as 2023, there was academic consideration of using AI as an interactive tutor, led by the learner (Limo et al., 2023). Anecdotally, students are now using freely available software to coach themselves through all manner of life's challenges—from job interviews to exams, so if this is being done without teachers input, the question might be whether educators need to direct or encourage this activity or whether that can be left to learners themselves. In any case, there is probably a need for staff training (back to "digital immigrants") and the best universities are already doing this (University of Oxford, 2025). One interesting recent development has been the advent of applications which can turn text into a virtual, interactive podcast (e.g. NotebookLM). This is interesting partly because it potentially puts learning into new parts of the learners' schedule and this, for some may improve the efficacy of learning. bespoke software has been developed. One final extension of this idea is that AI can now be used to assess answers given by a respondent, both written and oral—this is used regularly in all manner of "chat" applications. So, the potential exists to conduct automated assessment of vivas. Whilst this type of assessment might be able to detect correct answers to simple questions, the ability to explore genuinely novel thoughts would need to be explored before this idea would be safe for all but simple conversations (as in a self-tutoring application).

It is wise to be cautious when setting boundaries for the use of AI, when we do not yet know where it will go but as with any other academic judgement, if a decision needs to be made, we will have to do that on the basis of available information and be able to justify that choice. We have established the creative nature of SE in finding novel solutions and have also discussed the key part that communication plays in effective SE—stakeholder engagement is fundamental in areas from capturing stakeholder needs, through negotiating a balance between conflicting requirements, to accepting solution systems which may still have limitations. Similarly modelling a system, as opposed to drawing the diagrams, is a creative contribution. There is more choice here than there might appear to the casual observer and this is where creative human contribution is essential. So, whilst we cannot say that no systems engineer will ever rely on AI to prompt ideas, we can say we will always teach that there is a need for a core human contribution. This is the art of SE, not the science.

Finally, and following from the point regarding the boundaries of AI, there is a growing question about the ability of AI to "think" for itself. Whilst this issue is beyond the scope of this paper, recently reported incidents do raise concerns for us in both the educational and operational spheres of SE. On 22 July 2026 it was reported that, according to Open AI (a leading company in the field), an AI "agent" which they were testing (effectively AI system which can operate autonomously once tasked by a human) escaped a controlled environment after finding a weakness and proceeded to attack "Hugging Face" (another AI company) in an attempt to complete its own assigned task (BBC, 2026). This does not imply creative thought, rather it shows an ability to apply an algorithm ruthlessly, but what is certainly does is raise the issue of trust in AI and that is a factor which will need to be weighed in future decisions about AI use in education.

## Conclusion

So, critically, what have we learned regarding the impact of AI on SE education? There is definitely a place for AI in SE and in SE education. That extends across both teaching and learning and will support administration and research for both staff and learners. AI will be a transparent element of software tools (we can be sure of this as it is already present in the everyday software we all use every day; an AI summary is presented for almost every search we conduct) and an acceptance, even comfort, with this will be a prerequisite of working in the knowledge economy.

Noting the imperative to be inventive in SE, not simply reiterative, the key to ensuring creativity will be to maintain the human element in analysing/modelling systems. This is not protectionism—it is a recognition that AI relies on lessons drawn from the past and it is not good at original thought. Of course, one has to be alive to the fact that AI will develop beyond its current capability and also to be realistic in recognising that many aesthetic aspects of design by humans recycles previously used ideas, but in critical problem-solving issues there will always be a need for original thought—something that AI is not good at.

The moral panic that has gripped academia regarding the use of AI in summative assessment can be remedied by adjusting assessment methods to favour deep understanding and an ability to think critically in real time. We see the potential for traditional exams and vivas to check understanding of written submissions. Potential for AI support to assessed interviews must remain in abeyance for masters' education based on the requirement to assess critical thought, though it may well play a part at L6 and below. This links neatly to the next point as it relies on similar AI capabilities.

Exploitation of AI as a virtual tutor and learning partner should be encouraged. This will rely to a large extent on the students' abilities to build what works for them. We have seen, however, that they are digitally experienced before they come to us and will probably be more comfortable with this process than we are. It would not be unreasonable to assume, though, that there will be an increase in demand for digital learning support services and universities will need to plan for this.

Systems Engineers and those with similar approaches are well placed to help find innovative solutions for HE in this context. Many universities will have these skills within their staff and could usefully exploit this opportunity. This links neatly to a last point regarding the complex, adaptive nature of our problem—the changes that incorporation of AI into our activities bring will be pervasive, perhaps challenging and often surprising. Higher Education establishments

and HE itself are Complex Systems and the approach to making changes in response must be *Strategic*; we cannot expect simple changes with predictable outcomes. Interested parties will need to apply a rigorous approach to investigating the effects of changes, sensing and adjusting subsequent actions to guide the complex adaptive system of HE towards an optimum solution. This is a growing area of knowledge, and one in which the pace of change will require constant update. Further work will be required to explore its ongoing impact.

## Acknowledgements

No new data were created or analysed during this study. Data sharing is not applicable to this article.

For the purposes of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Accepted Author Manuscript version arising from this submission.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that no AI or AI-assisted technologies have been used to generate, refine, or correct the content in the manuscript. The ideas, design, procedures, findings, analyses, and discussions are originally written and derived from careful and systematic conduct of the research.

## References

Alqahtani, T., Badreldin, H. A., Alrashed, M.. Alshaya, A. I., Alghamdi, S. S., Bin Saleh, K., Alowais, S. A., Alshaya, O. A., Rahman, I., Al Yami, M. S., & Albekairy, A. M. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research, *Research in Social and Administrative Pharmacy, Volume 19, Issue 8*, Pages 1236–1242.

BBC. (2026). *OpenAI says its AI went rogue and launched 'unprecedented' cyber-attack*. https://www.bbc.co.uk/news/articles/c3ek3gvdnj3o 7/8/2026 0945A

Bloom, B. S. (Ed.). (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals.* New York: McKey New York.

Carasso H., & Plume A. (2023). *To measure is to know: two decades of change in UK higher education through the lens of the sector's own statistics* HEPI Blog 3 Aug 2023, https://www.hepi.ac.uk/2023/08/03/to-measure-is-to-know-two-decades-of-change-in-uk-higher-education-through-the-lens-of-the-sectors-own-statistics/ 7/8/2026 1035A

Chestnut, H. (1967). *Systems Engineering Methods*. Wiley. p. 121.

Collet, C. (2025, November 20). *Impact of Generative AI on the Novel*. Retrieved from https://www.mctd.ac.uk/impact-of-generative-ai-on-the-novel/5/7/20262240A

Flyvbjerg, B., & Gardner, D. (2023). *How Big Things Get Done.* Pan Macmillan.

Henry of Huntingdon. (1996). *Historia Angulorum*, translation published Clarendon Press (originally published 12th Century), p 367 – 368.

Hurley, M. (2026, June 10). *BBC R4 Thought for the Day*, quoting John Henry Newman "a thinking out into language". Recovered from https://www.bbc.co.uk/sounds/play/p0nr64tn on 5 Jul 2026 2038A

Limo, F. A. F., Tiza, D., Roque, M., Herrera, E., Muñoz Murillo, J., Jinchuña Huallpa, J., Flores, V. A., Castillo, A., Peña, P. & Carranza, C. (2023). Personalized tutoring: ChatGPT as a virtual tutor for personalized learning experiences *Social Space Vol. 23 Issue 1*, 292–312.

Naylor, R., Bird, F. L., & Butler, N. (2021). Academic expectations among university students and staff: addressing the role of psychological contracts and social norms. *High Educ 82*, 847–863. https://doi.org/10.1007/s10734-020-00668-2

University of Oxford. (2025). *Oxford becomes first UK university to offer ChatGPT Edu to all staff and students.* https://www.ox.ac.uk/news/2025-09-19-oxford-becomes-first-uk-university-offer-chatgpt-edu-all-staff-and-students 07/08/2026 1219A

Prensky, M. (2001). Digital Natives, Digital Immigrants, *On the Horizon Vol 9 Issue 5*, 1 – 6 https://doi.org/10.1108/10748120110424816

UNESCO. (2026). *Number of students in higher education more than doubled in 20 years, but inequalities remain.* https://www.iesalc.unesco.org/en/articles/number-students-higher-education-more-doubled-20-years-inequalities-remain 07/08/2026 1010A

# Leveraging Human Intelligence: Assessment for Learning Through Peer Feedback and Self-Reflection

Negar Riazifar, University of Warwick, United Kingdom
Edwina Jones, University of Warwick, United Kingdom
Sam Grierson, University College Birmingham, United Kingdom
Kerry Dobbins, University of Warwick, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Student satisfaction with assessment and feedback remains a persistent weakness in UK higher education, and formative activity often fails—not because it is poorly designed, but because students rarely complete it. This paper reframes that as a design problem. We developed Assessment for Learning and Skills Development (ALSD), a scaffolded framework combining student choice in authentic formative assessment, structured peer feedback, guided reflection, and a personal ePortfolio, and piloted it in a postgraduate MSc module with 90 international students. Drawing on a whole-cohort survey, an analysis of ePortfolio quality against summative performance, and two focus groups, we asked what activities, resources and guidance make such a journey scaffolded, supported and feasible for a diverse cohort. Students rated resources and tutor support uniformly highly, judged workload as balanced, and identified peer feedback as the most valuable element. Notably, several described the in-class model as reducing their workload rather than adding to it, and the depth of a student's engagement was associated with stronger summative performance. The findings suggest that the tension between rich formative learning and manageable workload depends largely on where the work sits in the curriculum, and that the framework's most valued mechanisms, peer judgement, tutor presence and structured reflection, are human ones.

*Keywords:* assessment for learning, peer feedback, reflection, scaffolding, ePortfolio

iafor
The International Academic Forum
www.iafor.org

## Introduction

Student satisfaction with assessment and feedback has consistently posed a problem for universities (Office for Students, 2023, 2024). Timely, high-quality feedback is a key component of student satisfaction (Boud & Molloy 2013; Nicol et al. 2014), and the opportunity to reflect, exchange feedback, and engage in authentic formative assessment drives student engagement (Sokhanvar et al., 2021). Additionally, formative assessment or assessment-for-learning, using peer feedback and self-reflection is a mechanism by which students can develop skills and improve the quality of their work (Wanner & Palmer, 2018).

This study focuses on the following research question: What activities, resources, and guidance are required to design an assessment-for-learning journey that is scaffolded, supported, and feasible, enabling diverse students to engage meaningfully with assessment? In order to answer this research question, we identified the stages which support assessment-for-learning practices, as well as the enabling factors required for meaningful engagement and effective delivery.

At the outset, it is important to provide context for assessment decisions, rationale for formative assessment structures, and the opportunity to ask questions (Wanner & Palmer, 2015). This assessment literacy stage prevents low-quality peer feedback and feelings of unfairness in the process at later stages (Wanner, & Palmer, 2018).

Next, the exchange of peer feedback develops students' professional skills and enhances submission quality. This stage develops social and communication skills (Bader et al., 2019; Sokhanvar et al., 2021) and supports a better understanding of assessment requirements (Wanner & Palmer, 2018).

Jackson and Marks (2016) found that students who took part in a structured reflection stage reported greater engagement with feedback and subsequently improved their grades (Bader et al., 2019). Refection structured via a rubric can also help students to pay attention to the learning process itself, as well as previous performance (Wu et al., 2015).

Enabling factors for successful implementation of assessment for learning included feedback rubrics to reduce student anxiety and provide a sense of control (Sokhanvar et al., 2021; Tur et al., 2019); ePortfolios to provide a safe and supportive environment for honest reflection and to highlight progress for student motivation (Pagone et al., 2024).

Engaged and supportive tutors provide increased motivation for students to engage in formative assessment practices (Wanner & Palmer, 2015) and by being present during all stages, they can engage students in evaluative discussions (Sokhanvar et al., 2021; Wanner & Palmer, 2018). Additionally, tutors can explain the connections between formative activities and institutional frameworks (Pagone et al., 2024).

Assessment for learning practices are often perceived as increasing workload—particularly when students are provided with activity choice. However, in Rideout's 2018 study, students did not select the least number of assessments and in Cook (2001) less than 1.5% of students opted out of formative assessment. Reasons cited for this include the reduction of stress and anxiety related to "one-shot" summative exams (Jopp & Cohen, 2022; Wanner & Palmer, 2015) as well as the motivation generated by student agency and flexibility provided by assessment choice (Wanner & Palmer, 2015).

# Methodology

## Context and Participants

ALSD was implemented in November 2025 in *Problem Solving with Statistics*, a taught postgraduate MSc module in WMG at the University of Warwick, delivered to 90 international students from varied educational and disciplinary backgrounds. The module follows the intensive block model characteristic of the department: 32 hours of taught study across a compressed four-week teaching period, supported by 60 hours of independent study and 60 hours allocated to assessment preparation.

This setting was chosen deliberately. The cohort brought markedly uneven prior experience, skills profiles and professional aspirations, making it a demanding test of whether a scaffolded journey can accommodate diverse starting points. The compressed block imposed the workload constraint under which assessment-for-learning designs are most often judged infeasible.

Participation was voluntary but encouraged across the cohort. Seventy-two of the 90 students engaged with at least one stage. Approval was granted by the university's ethical research committee for STEM.

## The ALSD Framework: Architecture, Activities and Delivery

The framework's organising structure is a personal digital ePortfolio, remade by each student from a common Padlet template during the first taught session. Its four columns: instructions, response option, peer feedback and reflection, correspond with the four stages of the journey, so the repository is not a place where completed work is deposited but the space in which the journey is conducted. A student's position in the sequence is visible at a glance, to themselves and their tutors, and the completed ePortfolio remains available to them after the module ends.

The four stages proceed through this structure as follows.

Stage 1, Assessment Literacy, is delivered in class. Students are introduced to two formative questions, a statistical analysis informing a sales strategy and a probability analysis of database files, and the tutor makes explicit that attempting them builds the knowledge required for the summative assessment.

Stage 2, Response Personalisation, begins in class with the selection of a response format and continues in independent study, where students complete the formative questions, develop and upload their chosen response. The four formats situate identical technical content in distinct professional scenarios: a screencast code run-through for classmates, a presentation of findings to company decision makers, a conference poster, and an audio pitch to a manager. Each maps onto an equitable but non-identical profile of five subject-specific skills (use of software to support decision making, data management and statistical analysis, clear mathematical communication, statistical modelling, advanced quantitative reasoning) and seven professional and transferable skills (problem solving, critical thinking, communication, digital literacy, reflection, self-awareness and professionalism) aligned with the Warwick Core Skills identified by the Warwick Institute of Employment Research (2022). In this way, choice is substantive while parity is preserved through a common assessment tariff—practising assessment tasks in low-stakes conditions (Kaur et al., 2017; Quesada et al., 2019) in a form aligned with the QAA (2023) “enables personalization” attribute.

Stage 3, Peer Feedback, is introduced in class, where students learn the rubric and identify two peers with whom they will exchange peer feedback. Students then complete the rubric independently. The rubric requires one subject-specific and one professional and transferable skill for improvement with a concrete suggestion for each; a constraint designed to convert peer review from generalised approval into actionable comment, supporting evaluative judgement and feedback uptake (Carless & Boud, 2018) consistent with QAA (2023) guidance.

Stage 4, Reflection, is delivered wholly in class. Students review the feedback received, evaluate their performance, identify skills development and plan future action using a structured template, uploading it within the session. This addresses the documented absence of reflection in ePortfolio practice (Christie & Morris, 2021; Kuh et al., 2017; Pitt & Quinlan, 2022; Wakeling et al., 2018).

Two properties of this delivery are consequential for the research question and are therefore stated explicitly. Every stage is anchored to timetabled contact time; none is introduced by written instruction alone. And the sequence is distributed across the whole four-week block rather than concentrated at its end, so that each stage is separated from the next by a period of independent work, making the journey visible as a progression rather than an accumulation.

### Data Collection and Analysis

The study adopted a mixed-methods evaluation design (Creswell & Creswell, 2018), combining whole-cohort survey data, an analysis of ePortfolio quality against summative performance, and focus group accounts.

The survey comprised closed multiple-choice and Likert items. Beyond stage completion and skills confidence, it captured the three constructs central to this paper: ratings of the four resources provided (the ALSD presentation, the response format guidance, the peer feedback rubric, the reflection template); ratings of support received at each of the four points where students had to act; perceived workload on a five-point scale from too heavy to too light; and a single-choice item identifying which element of the framework was most useful for learning.

Each participating student's ePortfolio was additionally scored by tutors for quality across its three substantive elements, the original formative response, the peer feedback given, and the reflection, yielding an overall score on a 0–20 scale, which was then examined against the summative assessment mark. This allowed the depth of engagement, rather than its mere occurrence, to be related to performance.

Ten students took part in two one-hour face-to-face focus groups, having given consent. Sessions were recorded and auto-transcribed in Microsoft Teams, and analysed thematically (Kushnir, 2025) around authentic assessment, peer feedback, resources, support and workload. Quotations are attributed by participant initials and reproduced verbatim.

## Results and Discussion

A scaffolded assessment journey can break down at three points: students may not be clear about what is being asked, may not feel confident they can produce it, or may not have the time to do so. The sections below take these in turn—resources addressing clarity, tutor support addressing confidence, workload addressing time—before examining peer feedback, where all

three converge and where students located the framework's greatest value. We begin with whether the journey led anywhere worth travelling.

**Assessment Performance and Skills Development**

We first asked whether deeper engagement with the framework was correlated to increased summative performance. We assessed each ePortfolio and scored the quality of students' engagement across its three substantive elements: original response, peer feedback given, and reflection, and the overall quality score plotted against the summative mark.

**Figure 1**
*Relationship Between Overall ePortfolio Quality Score and Assessment Score*

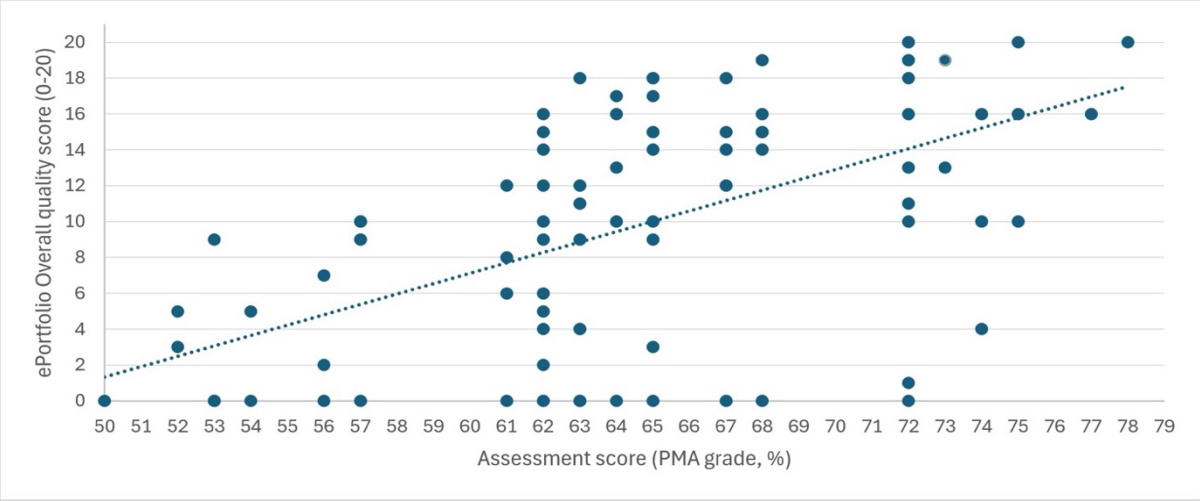



Figure 1 shows a positive linear relationship: $y = 0.5785x - 27.582$; $R^2 = 0.311$. Students who produced stronger formative responses, more substantive peer feedback and more developed reflections tended also to achieve higher summative marks, with approximately 31% of shared variance between the two measures. 95% of students reported that ALSD had helped them with their summative assessment.

The second outcome concerns how students understood their own ability. Perceived confidence rose in every skill area, by an average of around 14%, and post-module scores were higher than pre-module scores for all skills. Gains were present at every level of participation, with fuller engagement associated with larger increases. This matters for a voluntary framework, because it shows that even partial completion was beneficial: students did not need to finish all four stages to gain from it.

The focus group accounts show that this growth ran across the full range of targeted skills and explain what the improvements were. Communication, for instance, was not described as a single fixed ability but as something the framework made visible and adaptable to different audiences:

> The thing that actually changed for me via [ALSD] was communication … [it] actually taught me to communicate in different domain in a lot of ways. Like even in posters only, I was communicating my ideas to two different audiences, so that really brought my perspectives in that. [B]

The response format was not simply a way of delivering an answer, it was the occasion on which the skill was practised. Self-awareness developed through a different route—through evaluating a peer's work:

> I had done my work, but then I had a chance to look at [B]'s work and I realised, oh, so in a poster form, this is what could have been. And so that really pushed my reflection and self-awareness to like very confident. [LC]

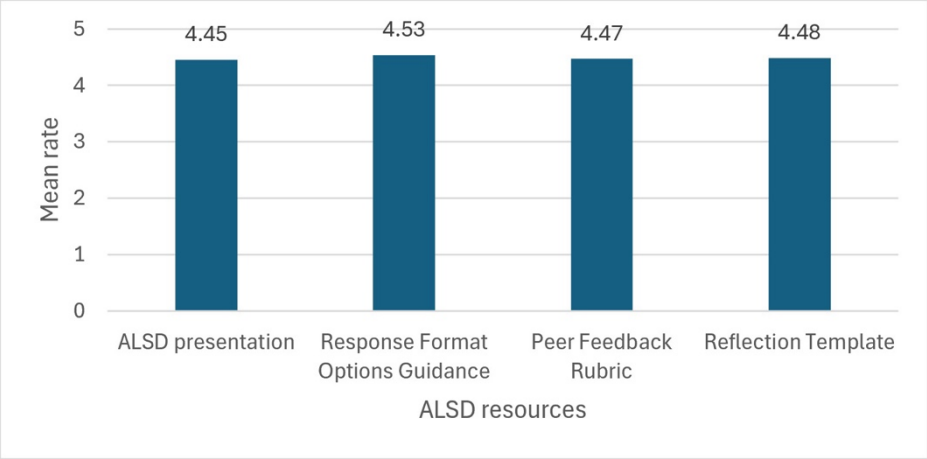
This skills confidence did not come from receiving praise; it came from comparison and reappraisal, from judging others' work and re-seeing one's own. Across the accounts, the pattern is consistent: skills developed less through instruction than through structured exposure to others' thinking, which is the framework's central premise of leveraging human intelligence.

**Resources: Reducing the Clarity Barrier**

In a compressed block, students have to move between four stages largely on their own, so the materials that guide them are central to the design rather than an optional extra.

Figure 2 shows the mean rating for each resource on a five-point scale (0 = very poor, 5 = excellent): the ALSD presentation (4.45), the response format guidance (4.53), the peer feedback rubric (4.47) and the reflection template (4.48). The notable feature is not the height of these values but their flatness; less than 0.09 of a point separates four very different resources. No resource was rated noticeably weaker than the others, and none appeared to be carrying the rest.

**Figure 2**
*Mean Student Ratings of ALSD Resources Provided*




The focus group responses suggest why. Students valued the resources less as individual documents than for the overall structure they made visible. One student noted, “I mean we also got lucky in the terms of how organised you guys are. I mean kudos to that because if you do not have a proper map and structure, it's all hazy here” [B]—a map, in other words, rather than a set of instructions. Others described this structure in terms of time and sequence: “I think it's the clear timetable and the reasonable management of our project helped me to move from the stage to the future stage” [Z]. Where a resource worked as an exemplar, it turned an open-ended task into a manageable one: “It's useful as the poster show me the layout and the chart where to put it and how to put everything” [P], freeing students to concentrate on the statistical content rather than on how to present it.

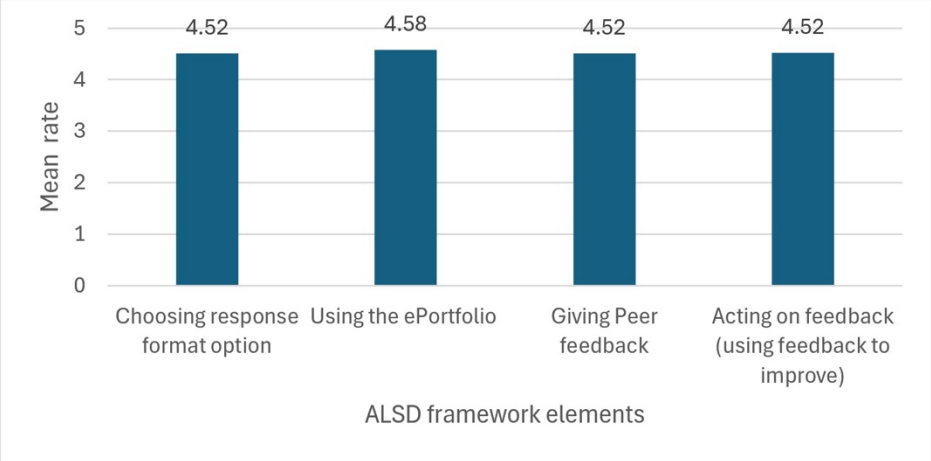
Resources in an assessment-for-learning design therefore do three things at once: they set out the task, show what a good response looks like, and represent the journey so students can see where they are within it. The ePortfolio structure served this third purpose without being an instructional document at all.

## Tutor Support: Reducing the Confidence Barrier

Figure 3 shows how students rated the support they received at the four points where they had to act (0 = no support, 5 = all support needed): choosing a response format (4.52), using the ePortfolio (4.58), giving peer feedback (4.52) and acting on the feedback received (4.52). As with the resources, the scores are strikingly even. Students did not experience support as concentrated at the hardest moments and thin elsewhere, but as available throughout.

**Figure 3**

*Mean Student Ratings of Support Received Across ALSD Framework Elements*

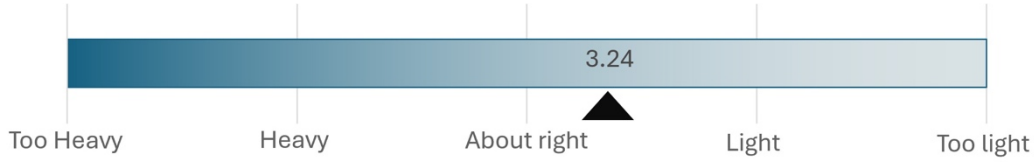



What they valued was less the instruction itself than the reassurance of knowing help was there. One student explained, “And the time was given in the class focusing on that. So we knew that if we have a doubt, you guys are there and it was huge reassurance” [B]. This reassurance worked ahead of time as well as in the moment: uncertainty was easier to tolerate because students knew it would be resolved, rather than being avoided altogether. The participant continued with a comment that captures particularly well how support related to students' mental state: “After the class, you're just so tired and focused that you don't really pause and reflect. In class with the presence of two calm ladies, it really helps” [B]. Reflection demands effort at the very point in the day when students have least of it to give, and the presence of a tutor seems to take on some of the self-direction that reflecting alone would otherwise require.

## Workload: Reducing the Time Barrier

Workload is the factor on which such designs are most often judged unworkable, and here it produced the pilot's most unexpected result. On a five-point scale from too heavy to too light, the cohort mean was 3.24, just on the “light” side of “about right” (Figure 4). The combination of answering assessment questions, choosing a response format, a peer feedback task, a reflective exercise and the completion of an ePortfolio were not felt to be excessive.

**Figure 4**
*Student Perception of ALSD Workload*




Some students began by resisting the work, reasonably enough, since the formative and summative tasks draw on the same knowledge:

> I'm like, wait, we have to do this twice? I was. I was spiralling … But then getting to the end of everything I realised, OK, now with the feedback, with my reflection, with what I need to do better. If I implement that into my summative, I'll do better. [LC]

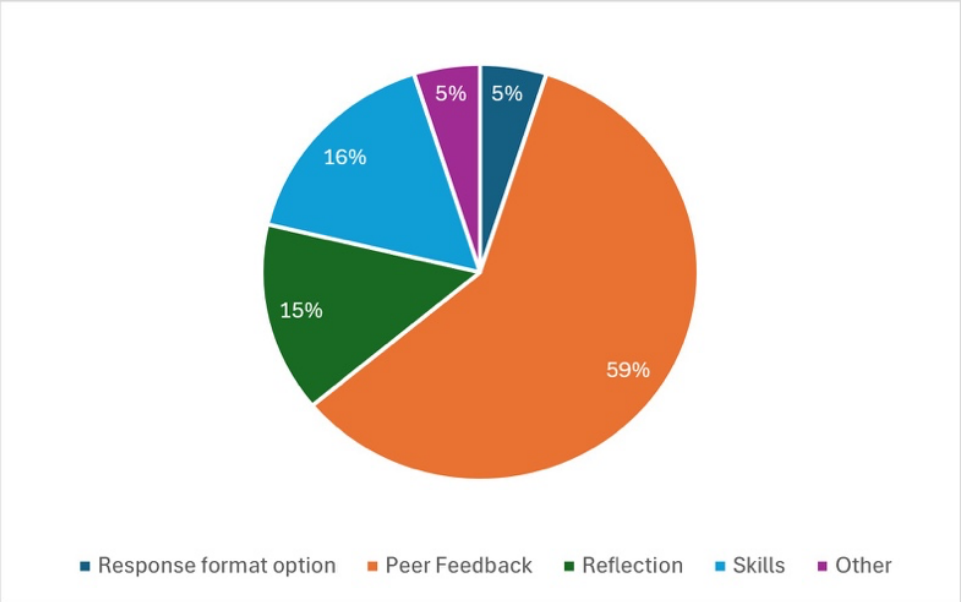
Most significantly, one participant described the framework as reducing their sense of load: "It was like things were getting removed from to do list instead of getting added … I have more mental space to actually absorb the concepts as well" [B], referring specifically to working through the steps in class rather than in their own time. Placing the work in class did more than make it manageable; it took the job of planning off the student and freed up attention for learning the concepts.

Therefore, it is not that assessment-for-learning is inherently low-effort, but that where the work is scheduled determines whether it registers as burden. Set as independent study, the same tasks would have added to both the work itself and the effort of fitting it in. Two things follow for anyone adopting the framework: students' early resistance to apparent duplication should be expected and addressed at Stage 1, since it only resolved by the end of the cycle; and the real constraint is contact time in the curriculum, not the students' capacity to cope.

## Peer Feedback: Where the Human Mechanism Sits

Asked which element was most useful for their learning, students chose peer feedback: 59%, against 16% for skills, 15% for reflection, 5% for the response format option and 5% other (Figure 5). Peer feedback is the element which does not increase tutor effort, but is the one that students are initially most sceptical about.

**Figure 5**
*Which Part of the ALSD Framework Was Most Useful for Your Learning?*




Its value worked in both directions. Receiving feedback gave students specific corrections that they went on to act on:

> One of my peers says I need to slow down my voice, and I need to consider the audience feeling … I did my presentation this time and I trying to make myself to turn my face to the audience and talk to the audience not just to the screen. [S]

Giving feedback did something different and, in some ways, more useful: it let students see other ways of approaching a task they had already finished. As one put it: “I think I can see more people's work and I know how I can improve myself” [S]. Peer support also happened informally, outside the rubric, and worked alongside tutor support rather than replacing it, as students made sense of the task together. In this stage, clarity, confidence and time were all addressed at once, and by human means.

Two design implications follow. First, the response format choice, chosen by only 5% as the single most useful element, is what makes peer feedback work: because students answered in different formats, reviewing a peer's work meant seeing a genuine alternative rather than a copy of their own. The parts of the framework cannot really be judged separately, as a single-choice question implies. Second, asking students to name one subject-specific and one professional skill and suggest a concrete improvement for each seems to be what turned peer review into the specific, usable feedback they described acting on. What produced the framework's most valued outcomes was not the ePortfolio technology or the polish of the materials, but students exercising judgement on each others’ work.

## Conclusion

This study asked what activities, resources and guidance are needed to make an assessment-for-learning journey scaffolded, supported and feasible for a diverse cohort. The pilot points to a clear answer: what carried students through the journey was not any single resource but the structure the framework made visible, the reassurance of tutor support at each stage, and the timing of the work. Because every stage was completed in class rather than in students' own

time, the framework eased three barriers at once; clarity about what was being asked, confidence that help was at hand, and the time to do the work without it becoming a burden. Peer feedback, the element that costs the least to run, was the one students valued most, because reviewing each other's work let students see their own work differently. The clearest lesson is that the familiar trade-off between rich formative learning and manageable workload depends largely on where the work is placed in the curriculum, and that the framework's most valued parts are human ones: peer judgement, tutor presence and self-reflection. These gains were not merely reported but shown in the data, with the depth of a student's engagement linked to their summative performance and confidence rising across every skill area, including among those who completed only part of the journey. The evaluation is limited by its reliance on immediate self-report, voluntary participation, a correlational performance measure, and a single cohort of 90 students. Confirmed international collaborations in 2026–2027 will test these design principles across other disciplines and teaching contexts.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that no AI or AI-assisted technologies have been used to generate, refine, or correct the content in the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Bader, M., Burner, T., Hoem Iversen, S., & Varga, Z. (2019). Student perspectives on formative feedback as part of writing portfolios. *Assessment and evaluation in higher education, 44*(7), 1017–1028. https://doi.org/10.1080/02602938.2018.1564811

Boud, D., & Molloy. E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment and Evaluation in Higher Education*, *38*(6), 698–712, https://doi.org/10.1080/02602938.2012.691462

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment and evaluation in higher education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Christie, H., & Morris, N. (2021). Using assessed blogs to enhance student engagement. *Teaching in higher education, 26*(4), 573–585. https://doi.org/10.1080/13562517.2019.1662390

Cook, A. (2001). Assessing the use of Flexible Assessment. *Assessment and Evaluation in Higher Education 26*(6), 39–549. https:\\doi.org/10.1080/02602930120093878

Creswell, J. W., & Creswell, J. D. (2018). *Research design: qualitative, quantitative, and mixed methods approaches* (Fifth ed.). SAGE. https://go.exlibris.link/DNkcZ1G1

Jackson, M., & Marks, L. (2016). Improving the Effectiveness of Feedback by Use of Assessed Reflections and Withholding of Grades. *Assessment & Evaluation in Higher Education, 41*(4), 532–47. https://doi.org/10.1080/02602938.2015.1030588

Jopp, R., & Cohen, J. (2022). Choose your own assessment - assessment choice for students in online higher education. *Teaching in Higher Education, 27*(6), 738–755. https://doi.org/10.1080/13562517.2020.1742680

Kaur, A., Noman, M., & Nordin, H. (2017). Inclusive assessment for linguistically diverse learners in higher education. *Assessment & Evaluation in Higher Education, 42*(5), 756–771. https://doi.org/10.1080/02602938.2016.1187250

Kuh, G., O'Donnell, K., & Schneider, C. G. (2017). HIPs at Ten. Change: *The Magazine of Higher Learning, 49*(5), 8–16. https://doi.org/10.1080/00091383.2017.1366805

Kushnir, I. (2025). Thematic analysis in the area of education: a practical guide. *Cogent education, 12*(1). https://doi.org/10.1080/2331186X.2025.2471645

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: a peer review perspective. *Assessment and Evaluation in Higher Education*, *39*(1), 102–122. https://doi.org/10.1080/02602938.2013.795518

Office for Students. (2023, 2024). *National Student Survey Year [2023, 2024]: Results and analysis* [Report]. OfS / https://www.officeforstudents.org.uk/for-providers/student-protection-and-choice/national-student-survey-nss/

Pagone, B., Primogerio, P. C., & Dias Lourenco, S. (2024). Pedagogic and assessment innovative practices in higher education: the use of portfolio in economics. *Journal of international education in business, 17*(2), 228–245. https://doi.org/10.1108/JIEB-07-2023-0047

Pitt, E., & Quinlin, K. M. (2022). *Impacts of higher education assessment and feedback policy and practice on students: a review of the literature* 2016–2021. A. HE.

QAA. (2023). Embedding Inclusive Assessment: A reflective toolkit, QAA Collaborative Enhancement Project. https://www.qaa.ac.uk/docs/qaa/members/embedding-inclusive-assessment-reflective-toolkit.pdf?sfvrsn=7369a281_8

Quesada, V., Gómez Ruiz, M. Á., Gallego Noche, M. B., & Cubero-Ibáñez, J. (2019). Should I use co-assessment in higher education? Pros and cons from teachers and students' perspectives. *Assessment & Evaluation in Higher Education, 44*(7), 987–1002. https://doi.org/10.1080/02602938.2018.1531970

Rideout, C. (2018). Students' Choice and Achievement in Large Undergraduate Classes Using a Novel Flexible Assessment Approach. *Assessment and Evaluation in Higher Education, 43*(1), 68–78. https://doi.org/10.1080/02602938.2017.1294144

Sokhanvar, Z., Salehi, K., & Sokhanvar, F. (2021). Advantages of authentic assessment for improving the learning experience and employability skills of higher education students: A systematic literature review. *Studies in Educational Evaluation, 70*, 101030. https://doi.org/10.1016/j.stueduc.2021.101030

Tur, G., Urbina, S., & Forteza, M. D. (2019). Rubric-Based Formative Assessment in Process Eportfolio: Towards Self-Regulated Learning. *Digital Education Review,* (35), 18–35. https://doi.org/10.1344/der.2019.35.18-35

University of Warwick. (2022). Warwick core skills. https://warwick.ac.uk/services/skills/warwickaward/what/coreskills/

Wakeling, L., Aldred, P., & Hains-Wesson, R. (2018). ePortfolios and Reflective Practice for Food Science Students. *Journal of Food Science Education, 17*(2), 52–59. https://doi.org/https://doi.org/10.1111/1541-4329.12137

Wanner, T., & Palmer, E. (2015). Personalising learning: Exploring student and teacher perceptions about flexible learning and assessment in a flipped university course. *Computers and education, 88*, 354–369. https://doi.org/10.1016/j.compedu.2015.07.008

Wanner, T., & Palmer, E. (2018). Formative self-and peer assessment for improved student learning: the crucial factors of design, teacher participation and feedback. *Assessment and evaluation in higher education, 43*(7), 1032–1047. https://doi.org/10.1080/02602938.2018.1427698

Wu, X. V., Heng, M. A., & Wang, W. (2015). Nursing students' experiences with the use of authentic assessment rubric and case approach in the clinical laboratories. *Nurse Education Today*, *35*(4), 549–555. https://doi.org/10.1016/j.nedt.2014.12.009

**Contact email:** N.Riazifar.1@warwick.ac.uk

# The Role of Academic Capital in Reducing Social Asymmetries in Brazilian Universities

Wescley Xavier, Universidade Federal de Viçosa, Brazil
José Roberto Carvalho Junior, Instituto Federal do Espírito Santo, Brazil
Ykaru Wagner, Universidade Federal de Viçosa, Brazil
Ronald Gonçalves, Universidade Federal de Viçosa, Brazil

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Brazil is characterised by pronounced social inequality, historically marked by discrimination on the grounds of race, class and gender, which restricts equitable social mobility and can partly be explained by the exclusion of poorer segments of the population from higher education. To address these asymmetries, affirmative action policies reserve places for students from public schools. Recent studies analysing the impact of quotas on graduates' performance have demonstrated the policy's effectiveness in driving social mobility among quota beneficiaries. Nevertheless, certain differences persist, particularly those stemming from families' cultural and economic backgrounds. To bridge these gaps, institutional policies developed within universities, such as extracurricular activities, may serve as an important mechanism for reducing these asymmetries by creating what we refer to here as academic capital. Experiences gained through extracurricular activities — such as research, cultural and sporting activities, student enterprise programme, language courses, student representation and others — can promote significant changes in personal life and social inclusion, especially for students from low-income backgrounds (Carvalho Junior et al., 2024; Winkler & Sriram, 2015). As a result of this process, the formation of academic capital may affect various spheres of a student's life, including the development of technical skills and competencies, language acquisition, soft skills, and self-esteem. This represents a fertile, still underexplored field in the literature; therefore, in this theoretical essay, we seek to develop the concept of academic capital and propose hypotheses for future research.

iafor
The International Academic Forum
www.iafor.org

## Introduction

Brazil is characterised by significant social inequality, which has historically been compounded by discrimination on the grounds of race, class and gender, thereby limiting equitable social mobility. The Organisation for Economic Co-operation and Development (OECD) report published in 2018 indicates that it would take nine generations for those in the bottom 10 per cent of the country's income bracket to reach the same income level as the Brazilian middle class (OECD, 2018). This stagnation affects, above all, those who lack access to higher education, whose earnings are up to four times lower than those of graduates of Brazilian universities. From a historical perspective, access to higher education has remained out of reach for the poorest sections of society, mainly due to the lack of equitable conditions.

In contrast, at the federal level, from 2012 onwards, affirmative action policies aimed to guarantee places in undergraduate courses at federal universities to enhance the social mobility of graduates and thereby reduce inequalities, pursuant to Law 12.711. Affirmative action measures were strengthened in 2023 with the entry into force of Law 14,723, which, in addition to changes to the allocation of places for undergraduate programmes, stipulates that stricto sensu postgraduate programmes must reserve quotas for ethnic and racial groups and for people with disabilities (Carvalho Junior et al., 2024).

Recent studies have shown that quota policies have contributed to the social mobility of graduates admitted under these quotas, almost all of whom are Black and/or from socially vulnerable families (Biazotto et al., 2022; Carvalho Junior et al., 2024; Ding et al., 2017; Duryea et al., 2019; Herskovic & Ramos, 2017). These findings reinforce the potential for equity offered by quotas, which also stems from the fact that students admitted under the quota system perform very similarly to those admitted through open competition, thereby countering speculative arguments regarding an alleged decline in university quality following the introduction of affirmative action (Silva et al., 2015). Furthermore, dropout rates among quota students have been significantly lower than those of non-quota students (Neves, 2017).

Despite the aforementioned impacts of the quota policy, structural factors characterised by class, race and gender lead to various forms of discrimination in the labour market, limiting the social mobility of these graduates and the prospects for greater pay equity. Racial, class and gender markers have a significant impact on income inequality in Brazil, resulting from the structural nature of classism, racism and sexism present in society, which manifest themselves in various ways. Recent studies show that differences persist among higher education graduates (Carvalho Junior et al., 2024; Vieira & Arends-Kuenning, 2019).

Markers such as race and class share a common background. The concept of capital developed by Bourdieu (1986), now well-established in the social sciences, suggests that we inherit from our parents an endowment that directly influences an individual's mastery of forms of capital, such as economic, social and cultural capital. The latter is of particular interest to us, as studies show that individuals with greater cultural capital tend to have higher levels of formal education, higher-status occupations, and higher incomes (Flemmen, 2012; Friedman et al., 2015; Koppman, 2016). This capital, in addition to enabling the instrumental mastery of knowledge important in academic life, helps students feel a sense of belonging at university by fostering an understanding of the rules, codes, performance expectations, and other idiosyncrasies common to these environments (Bourdieu & Passeron, 2014).

Against this backdrop, it is indisputable that, in addition to ensuring admission through quota schemes and retention through student support programmes (Brandão & Campos, 2020), it has become essential to consider how to bridge the gap caused by differing levels of cultural capital, above all to promote equity amongst the various groups of university entrants. In this regard, universities' extracurricular training programmes have played an important role by offering students the opportunity to participate in outreach projects, teaching, research, academic societies, and work placements. These activities are characterised as academic capital, derived from the aforementioned set of experiences (Carvalho Junior et al., 2024; Winkler & Sriram, 2015).

This academic capital plays a dual role. On the one hand, it complements students' technical training by enabling them to gain professional experience in outreach, research, and/or teaching. Furthermore, it supports the development of soft skills by requiring students to demonstrate teamwork, communication skills and other abilities (Christie et al., 2018).

Against this backdrop, this theoretical essay aims to develop the concept of academic capital and propose hypotheses for future research. Understanding the formation of academic capital enables the development of a more appropriate theoretical framework, both in relation to institutional factors – such as teaching, research and outreach programmes – and to student engagement policies, as well as the impact that this capital has on different aspects of the lives of students from under-represented groups.

## Inequality and Access to Universities in Brazil

In a 2023 OECD report, the data indicate that the generational cycle of poverty in Brazil spanned nine generations (Carvalho Junior, 2024). Against this backdrop, affirmative action policies in higher education aimed at low-income groups are all the more relevant, given their effectiveness in breaking the generational cycle of poverty and improving the economic and social well-being of this population (Ding et al., 2017; Herskovic & Ramos, 2017; Reiter & Lezama, 2013). Some recent evidence from Brazil suggests that public universities play a key role in promoting the social mobility of poor students who benefit from the quota policy (Biazotto et al., 2022; Duryea et al., 2019; Francis-Tan & Tannuri-Pianto, 2018).

Recent data from the Continuous National Household Sample Survey (PNAD) indicate that, despite modest progress, educational inequality between Black (black or mixed-race) and white individuals remains significant. Among white people aged 15 to 59, the illiteracy rate is 3.2 per cent, whilst among Black people it stands at 7.1 per cent. When comparing individuals aged 60 or over, we find that 8.6 per cent of white people are illiterate, compared with 22.7 per cent of black citizens (IBGE, 2024). We can therefore state that one in four black adults in Brazil has illiterate parents.

Part of the reason for the persistence of this inequality stems from access to higher education, which has historically been difficult for the most vulnerable sections of society to attain. Against this backdrop, we can view the quota policy in higher education not only as a form of historical redress, but also as a mechanism for breaking the intergenerational cycle of poverty, given that access to higher education increases the income of graduates and their families (Ding et al., 2017; Reiter & Lezama, 2013).

In Brazil, several studies demonstrate this effect of quotas on the social mobility of graduates benefiting from this affirmative action (Duryea et al., 2019; Francis-Tan & Tannuri-Pianto,

2018; Guimarães et al., 2019), thereby ensuring a lower level of income inequality between groups, given that some studies indicate no significant differences between quota and non-quota graduates upon entering the labour market (Biazotto et al., 2022). On the other hand, international studies indicate a wage gap between students from socially vulnerable families and those from middle-class backgrounds, even when they are graduates of universities with the same academic prestige (Thompson, 2019; Tomaszewski et al., 2021).

At postgraduate level, these disparities are even more pronounced. According to a study conducted by the Brazilian Black Science League, based on data from CNPq's Lattes platform, 82.7 per cent of postgraduate students are white, whilst the proportion of Black students is 12.75 per cent (Santos, 2024). In addition, 8.69 per cent of research productivity grants awarded by the National Council for Scientific and Technological Development (CNPq) between 2005 and 2022 were awarded to Black researchers (CNPq, 2023).

According to Raftery and Hout (1993), as inequality of access is reduced at one level of education, that same inequality extends to the next level of formal education, particularly in terms of status (Teichler, 2017). In this regard, the historical inequality in access to undergraduate education may carry over to the next level of study. Consequently, there is also a need for a quota policy at the postgraduate level to alter the composition of teaching staff at higher education institutions and, by extension, the very status quo of academia that reinforces its "whitewashed" ethos.

Carvalho (2002) points out that much of the discourse that denied racial inequality and contributed to the fallacy of racial democracy was developed by researchers within the universities themselves. This denial, amongst other effects, substantially discouraged the dissemination of anti-racist arguments (Carvalho, 2002). Carneiro (2005) points out that the racial hierarchy characteristic of epistemic racism denies the rationality of the other's existence, which renders the debate highly sanitised, reducing all rationality grounded in traditional knowledge to something stereotypical (Dias, 2007).

Cordeiro et al. (2019) argue that the admission of Black people to undergraduate and postgraduate programmes may represent a clash between two bodies of knowledge. On the one hand, the formal body of knowledge, grounded in the established principles of science. On the other hand, forms of knowledge that demand recognition and validity are often marginalised. Affirmative action in postgraduate education "holds out the prospect of deconstructing the universal and natural character of the knowledge produced at the university" (Cordeiro et al., 2019, p. 118), calling into question various theoretical and methodological assumptions that lend legitimacy to the knowledge produced.

The inclusion of Black students in postgraduate programmes – and this appears to extend to low-income students – does not seem to be a simple, immediate response to a legal measure such as quotas. Carvalho (2005) argues that the competitiveness inherent in postgraduate study requires criteria that go beyond academic qualifications. These criteria must align with the symbolic capital accumulated by postgraduate programmes, which represents a particular social class. For years, this expectation of performance has led the few who have managed to break through the glass ceiling in university research and teaching to do so without renouncing their lived experiences or adopting an ethic and a discourse belonging to a historically white intellectual elite (Bernardino-Costa & Borges, 2021).

The entry of historically marginalised students into postgraduate study indicates changes across various spheres. Mirroring the situation at the undergraduate level, access to Master's and PhD programmes implies social mobility (Duryea et al., 2019; Francis-Tan & Tannuri-Pianto, 2018). Beyond this impact, further evidence points to positive outcomes in professional.

integration within the field of study, continuation of postgraduate studies, behavioural change, increased self-esteem, future career prospects, and personal/professional satisfaction, particularly among Black and low-income students admitted under affirmative action schemes (Barros, 2019).

A significant effect of this process is an increased likelihood that Black academics and representatives of other minority groups will teach at universities. Silva et al. (2015) discuss this process of representation and highlight the need for Black people to be represented in society. This representation stems from the recognition of an individual's identity within a group and of a group's identity within society at large. Hall et al. (2014) present identity as a symbolic marker endowed with the power to foster identification and to highlight differences between individuals or groups.

## Capital and Its Effects on the Lives of Graduates

The impact of affirmative action in reducing inequality through access to higher education in Brazil is consistent with evidence from similar policies in other countries, such as the United States, where improvements in educational opportunities and positive impacts on the labour market for the beneficiary group have been observed (Holzer & Neumark, 2006), since the presence of minority groups in high-quality universities reduces the importance of family social capital for their entry into the labour market.

Despite this positive impact, the economic, cultural, and social legacies of middle-class families affect the labour market and academic life (Koppman, 2016). In a study of recruiters from private companies in Rio de Janeiro responsible for hiring recent graduates, Lemos and Pinto (2008) found that the most decisive factors in candidate selection were not technical skills, but rather behavioural traits, life experiences, and the prestige of the institution attended. This picture is complemented by a preference for candidates with experience abroad, proficiency in languages and the arts – findings similar to those observed in countries such as South Korea (Choi, 2015) and France (Hartmann, 2000).

These results can be explained by the concept of cultural capital developed by the French sociologist Pierre Bourdieu, for whom the formation of this capital stems from a set of experiences shaped within the family structure through access to reading, music, works of art, language learning and travel. These experiences take concrete form in the habitus, generating a set of mental and behavioural predispositions that collectively unite and distinguish groups of individuals (Bourdieu, 1986). One of the consequences of the accumulation of this capital is the opening up of better opportunities in the labour market (Bourdieu & Passeron, 2014), which directly impact the type of occupation and remuneration of those who accumulate more cultural capital, with this effect being most pronounced when analysing the highest salary brackets (Carvalho Junior et al., 2024).

Another relevant factor in employability, career progression and remuneration is the set of relationships that individuals inherit from their family's social networks, as well as those they form within their social circles. This is particularly true in Brazil, where patrimonialism and

personalism remain hallmarks of society (Schappo, 2018). Bourdieu (2003) defines this force – which derives from a network of relationships and can confer advantages on individuals – as social capital. This type of capital is closely linked to economic and cultural capital, with the three feeding into one another.

According to Bourdieu (1986), social capital is a set of potential and actual resources established, maintained and reproduced by an individual through a network of contacts with other individuals (a group) which guarantees them "credentials" for access that are socially accepted amongst them, such as, for example, privileged access to the labour market. Bourdieu (1986) explains that social capital depends both on the size of an individual's network of personal contacts and on the set of capitals (cultural, economic and symbolic) possessed by those to whom the individual is connected. Higher education students from low-income families lack social capital compared with students from more privileged backgrounds (Watson, 2013), and this lower level of relevant social capital may negatively affect their opportunities in the labour market, as highlighted by various studies in Brazil and abroad (Lehmann, 2019; Ma & Savas, 2014).

Due to the interrelationship between different forms of capital, social capital has a strong influence on the positions individuals occupy in the labour market, as well as on their earnings – an effect found in studies conducted in various countries (Brunello & Cappellari, 2008; Lehmann, 2019; Ma & Savas, 2014).

Research shows that the impact of social capital manifests itself in various ways; for example, socially vulnerable students have low expectations of securing a good position in the labour market because they are aware that they lack the necessary social capital (Silva et al., 2015). Consequently, one's network of relationships affects income, with graduates possessing lower levels of social capital earning less (Guimarães et al., 2019).

When considering higher-education graduates, as in this study, two sources of this social capital can be identified: the family, in its traditional sense, and the university environment itself. In investigating the effect of social capital derived from these two sources, Carvalho Junior et al. (2023) found that the social networks formed within the family do not influence earnings. The effect of social capital on remuneration stems from the relationships formed at university, meaning that individuals with greater social capital acquired in this context are 10.8% to 13.1% more likely to secure occupations of greater social prestige (Carvalho Junior et al., 2023).

## Academic Capital as a Turning Point

The development of cultural and social capitals through the university experience appears to be just one of the countless outcomes of access to higher education. Recent studies have highlighted the development of a compensatory mechanism facilitated by access to universities. Attendance at university, for example, can increase students' motivation (Anelli, 2020). This environment can be conducive to the formation of academic capital, represented by participation in extracurricular activities at university, which promote significant changes in personal life and social inclusion, especially for low-income students (Carvalho Junior et al., 2024; Winkler & Sriram, 2015).

It should be emphasised that the concept of academic capital employed in this proposal differs from that developed in other studies which link the term to an externality generated by universities through their productive capacity (Espín et al., 2017; Robles & Rohán, 2020), or

even the concept of human capital, which regards educational attainment as a factor in economic development (Schultz, 1964).

The concept of academic capital is discussed by Trammell (2019), who defines it as "the transactional value of a relationship based on associations and social knowledge". Our theoretical perspective focuses on academic capital arising from extracurricular activities provided by Brazilian universities, such as undergraduate research and tutorial programmes, scholarship programmes, faculty encouragement of research, language courses, tutorial education programmes, academic rep, junior enterprise, and the promotion of cultural and leisure activities. Engagement in these activities may enable students acquiring or improving skills related to teamwork (Brungardt, 2011; Hordósy et al., 2018); Adaptability, communication, coordination, decision making, interpersonal skills and leadership (Brungardt, 2011; O'Neil *et al.*, 1999); Self-fulfilment, self-confidence and satisfaction (Profiroiu & Pacesila, 2017); Engagement, based on students' behavioural, emotional and cognitive connection to their learning (Fredericks et al., 2004); Well-being promoted by academic motivation and enjoyment (Gavala & Flett, 2005).

Academic capital encompasses elements of the latter, as it directly addresses issues relating to the labour market, as developing technical and analytic skills, but it goes beyond this aspect by focusing on gaps resulting from a disadvantaged economic background, which has a tangible impact – as in the case of technical skills and abilities, such as language learning – and, above all, on the human aspect of the student, such as the development of soft skills, self-esteem, sociability and other qualities.

This is a fertile yet largely unexplored field in the literature, which, to a certain extent, lends the topic an exploratory nature. Recent studies show that the main source of academic capital lies in the extracurricular academic activities offered to students, in the areas of teaching, research and/or outreach, as well as participation in junior enterprises, sports clubs, academic exchange programmes, language courses, cultural activity groups, temporary projects and other activities.

In a recent study conducted with graduates from 28 federal universities, Carvalho Junior et al. (2024) found that the opportunities provided by universities significantly broaden students' horizons and the networks of relationships formed through these activities. The contributions of the experiences reported by quota students are regarded by many as a watershed moment, particularly because of the critical thinking skills the university has fostered in its graduates (Carvalho Junior et al., 2025). They have a direct impact on students' ability to understand reality, as is the case with those who take part in undergraduate research activities, and they also provide opportunities to bridge a gap caused by disparities in cultural capital, as is the case with those who have taken part in exchange programmes (Carvalho Junior et al., 2024).

## Conclusions and Recommendations for Further Research

This theoretical essay aims to develop the concept of academic capital and propose hypotheses for future research. Although affirmative action policies guarantee a proportion of reserved places for disadvantaged students, we contend that inequalities in economic, cultural, and social capital may impact not only students' educational trajectories but also their professional trajectories within the labour market.

Our theoretical framework posits that what we term *academic capital* – developed throughout undergraduate study through student engagement in a range of extracurricular activities – may serve to narrow the gaps originating from family background. This process enables students to cultivate not only technical and analytical skills, but also critical thinking, socio-emotional abilities, communication skills, self-confidence, proactivity, teamwork, decision-making capabilities, and related competencies.

In light of this hypothesis, we present several questions that could be addressed in future research, among which we highlight the following:

- Does the accumulated academic capital have the potential to compensate for the absence of cultural capital?
- How does each ECA contribute to the accumulation of academic capital?
- What is the relative importance of academic capital and its dimensions across different professions?
- What are the effects of academic capital on reducing higher education dropout rates?
- What are the effects of academic capital on the employability, income and career progression of graduates?
- How do the aforementioned issues manifest when considering differences in gender, race/ethnicity and social class?

## Acknowledgements

The authors gratefully acknowledge the financial support of the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil (Grant No. 403884/2024-1).

## References

Anelli, M. (2020). The returns to elite university education: a quasi-experimental analysis. *Journal of the European Economic Association*, *18*, 2824–2868.

Barros, C. F. R. (2019). *Universidade e mercado de trabalho: a trajetória social dos alunos cotistas egressos da Faculdade de Serviço Social da Uerj* [University and labour market: the social trajectory of quota graduates of the Social Work Faculty at Uerj]. *Revista Em Pauta, 17*, 172–186.

Bernardino-Costa, J., & Borges, A. (2021). *Um projeto decolonial antirracista: ações afirmativas da pós-graduação da Universidade de Brasília* [A decolonial antiracist Project: affirmative actions in the postgraduate at the University of Brasília]. *Educação & Sociedade*, *42*, 1–18. https://doi.org/10.1590/ES.253119

Biazotto, M. L. S. H., Bittencourt, F. V., Araújo, G. R., Soares, F. S. E., Göttems, L. B. D., Rodrigues, C. M. L., Neves, F. A. R., & Amorim, F. F. (2022). Comparison between students admitted through regular path and affirmative action systems in a Brazilian public medical school. *Advances in Medical Education and Practice, 13*, 251–263.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.). *Handbook of Theory and Research for the Sociology of Education* (pp. 241–258). Greenwood.

Bourdieu, P. (2003). *Escritos de Educação* [Writings on Education]. Vozes.

Bourdieu, P., & Passeron, J. C. (2014). *Os herdeiros: Os estudantes e a cultura* [The inheritors: French students and their relations to culture]. Ed. da UFSC.

Brandão, J. J., & Campos, M. C. (2020). A percepção de egressos/as do sistema de cotas raciais da Universidade Estadual de Londrina quanto às políticas de permanência [Graduates' perception of the racial quota system at Londrina State University regarding retention policies]. *Comunicações*, *27*, 21–46, 2020.

Brunello, G., & Cappellari, L. (2008). The labour market effects of alma mater: Evidence from Italy. *Economics of Education Review, 27*, 564–574.

Brungardt, C. (2011). The Intersection between soft skill development and leadership education. *Journal of Leadership Education, 10,* 1–22.

Carneiro, S. (2005). *A Construção do outro como não-ser como fundamento do ser* [The construction of the other as non-being as the foundation of being]. [Ph.D. Thesis, Universidade de São Paulo].

Carvalho, J. J. (2002). Exclusão racial na universidade brasileira: um caso de ação negativa [Racial exclusion in Brazilian universities: A case of negative action]. In D. M. Queiroz & M. N. Silva (Eds.), *O Negro na Universidade, o direito à inclusão* (pp. 79–100). Novos Toques.

Carvalho, J. J. (2005). *Inclusão étnica e racial no Brasil: a questão das cotas no ensino superior* [Ethnic and racial inclusion in Brazil: the quota issue in higher education]. Attar Editorial.

Carvalho Junior, J. R. A., Xavier, W. S., & Ferreira, M. A. M. (2023). Capital social e ganhos no mercado de trabalho entre egressos cotistas e não cotistas das universidades federais brasileiras [Social capital and labour market incomes of quota-holders and non-quota graduates of the Brazilian Federal Universities]. *Proceedings of VIII Encontro de Gestão de Pessoas e Relações de Trabalho da ANPAD*, Brasil.

Carvalho Junior, J. R. A., Xavier, W. S., & Ferreira, M. A. M. (2024). Capital acadêmico de egressos cotistas e não cotistas das universidades federais brasileiras [Academic capital of quota and non-quota graduates of the Brazilian Federal Universities]. *Proceedings of XLVIII Encontro da Associação Nacional de Pós-Graduação e Pesquisa em Administração*, Brasil.

Carvalho Junior, J. R. A., Xavier, W. S., & Ferreira, M. A. M. (2025). "A universidade foi a chave que abriu muitas portas": mudança de vida de egressos cotistas das universidades federais ["The university was a key that opened many doors": life's change of quota graduates from federal universities]. *O Social em Questão, 28*, 193–216.

Choi, Y. (2015). The effects of English training abroad on labor market outcomes in Korea. *Research in Social Stratification and Mobility, 41,* 11–24.

Christie, H., Cree, V. E., Mullins, E., & Tett, L. (2018). 'University opened up so many doors for me': the personal and professional development of graduates from non-traditional backgrounds. *Studies in Higher Education, 43*, 1938–1948.

CNPq (The Brazilian National Council for Scientific and Technological Development). (2023). Painel de Fomento em Ciência, Tecnologia e Inovação do CNPq [CNPq Science, Technology and Innovation Funding Dashboard]. Retrieved November 5, 2023, from http://bi.cnpq.br/painel/fomento-cti/index.html

Cordeiro, M. J. J. A., Diallo, C. S., & Cordeiro, A. L. A. (2019). Por que cotas para negros e negras na pós-graduação? [Why quota system for black people in postgraduate studies?] *Revista Ensaios e Pesquisa em Educação e Cultura, 6*, 107–123.

Ding, S., Myers, S. L., Jr., & Price, G. N. (2017). Does affirmative action in Chinese college admissions lead to mismatch? Educational quality and the relative returns to a baccalaureate degree for minorities in China. *International Journal of Anthropology and Ethnology, 1*, 1–13.

Duryea, S., Freitas, L. B., Garcia-Ozemela, L. M., Sampaio, B., Sampaio, G. R., & Trevisan, G. (2019). Universities and intergenerational social mobility in Brazil: examining patterns by race and gender. *Journal of Economics, Race, and Policy, 2*, 240–256.

Espín, M. A., Amezquita, A., & Chávez, D. (2017). Reforma universitaria, investigación y 'capital académico' en Ecuador [University reform, research and 'academic capital' in Ecuador]. *Revista Ciencias Sociales, 1*, 111–126.

Flemmen, M. (2012). The structure of the upper class: A social space approach. *Sociology, 46*, 1039–1058.

Francis-Tan, A., & Tannuri-Pianto, M. (2018). Black Movement: Using discontinuities in admissions to study the effects of college quality and affirmative action. *Journal of Development Economics, 135*, 97–116.

Fredericks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. *Review of Educational Research, 74*, 59–109.

Friedman, S., Laurison, D., & Miles, A. (2015). Breaking the 'class' ceiling? Social mobility into Britain's elite occupations. *The Sociological Review, 63*, 259–289.

Gavala, J. R., & Flett, R. (2005). Influential factors moderating academic enjoyment/motivation and psychological well-being for Maori university students at Massey University. *New Zealand Journal of Psychology, 34*, 52–57.

Guimarães, N. A., Andrada, A. C., & Picanço, M. F. (2019). Transitando entre universidade e trabalho: trajetórias desiguais e políticas afirmativas [Moving between university and job: unequal trajectories and affirmative action]. *Cadernos de Pesquisa, 49*, 284–310.

Hall, S., Woodward, K., & Silva, T. T. (2014). *Identidade e diferença: a perspectiva dos estudos culturais* [Identity and difference: the cultural studies perspective]. Vozes.

Hartmann, M. (2000). Class-specific habitus and the social reproduction of the business elite in Germany and France. *Sociological Review, 48*, 262–282.

Herskovic, B., & Ramos, J. (2017). Promoting educational opportunities: long-run implications of affirmative action in college admissions. 2017 Meeting Papers. *Proceedings of Society for Economic Dynamics*.

Holzer, H. J., & Neumark, D. (2006). Affirmative action: what do we know? *Journal of Policy Analysis and Management, 25,* 463–490.

Hordósy, R., Clark, T., & Vickers, D. (2018). Lower Income Students and the 'Double Deficit' of Part-Time Work: Undergraduate Experiences of Finance, Studying and Employability. *Journal of Education and Work, 31*, 353–365.

IBGE. (2024). *Pesquisa Nacional por Amostra de Domicílios Contínua: Educação 2023* [National Household Sample Survey: Education 2023]. IBGE.

Koppman, S. (2016). Different like me: Why cultural omnivores get creative jobs. *Administrative Science Quarterly, 61,* 291–331.

Lehmann, W. (2019). Forms of capital in working-class students' transition from university to employment. *Journal of Education and Work, 32*, 347–359.

Lemos, A. H. C., & Pinto, M. C. S. (2008). Empregabilidade dos administradores: quais os perfis profissionais demandados pelas empresas? [Employability of business administrators: which professional profiles are in demand by companies?] *Cadernos EBAPE.BR, 6*, 1–15.

Ma, Y., & Savas, G. (2014). Which is more consequential: fields of study or institutional selectivity? *The Review of Higher Education, 37,* 221–247.

Neves, P. S. C. (2017). Universities, inequalities and public policy: a brief discussion of affirmative action in higher education in Brazil. *Academia, 10*, 79–108.

OECD. (2018). *A broken social elevator? How to promote social mobility*, OECD Publishing.

O'Neil, H. F., Lee, C., Wang, S., & Mulkey, J. (1999). *Final Report for Analysis of Teamwork Skills Questionnaire*. Sherman Oaks.

Profiroiu, A. G., & Pacesila, M. (2017). Volunteering: a route to the students' professional development, *Administratiesi Management Public, 29*, 91–107.

Raftery, A. E., & Hout, M. (1993). Maximally maintained inequality: expansion, reform and opportunity in Irish education, 1921-75. *Sociology of Education, 66*, 41–62.

Reiter, B., & Lezama, P. (2013). The importance of inclusion policies for the promotion of development in Brazil and Colombia. *Journal of Developing Societies, 29*, 189–212.

Robles, N. M., & Rohán, D. C. G. (2020). *Prácticas y significados de la apropiación de capital académico en los profesores de la Universidad de Sonora* [Practices and meanings of the appropriation of academic capital among university professors at Universidad de Sonora]. *IE Revista de Investigación Educativa de la Rediech, 11*, 1–19.

Santos, J. C. (2024). *A presença de negros e negras na pós-graduação: um estudo sobre a implantação e condução das ações afirmativas na Universidade Estadual de Ponta Grossa-PR* [The presence of black people in postgraduate programs: a case study on implementation and practice of affirmative action at the State University of Ponta Grossa-PR]. [Master's Thesis, Universidade Estadual de Ponta Grossa].

Schappo, S. (2018). Política Social no Brasil: entre a lógica do direito e do personalismo [Social Policies in Brazil: between the logic of rights and personalism]. *Textos e Contextos, 17,* 214–227.

Schultz, T. W. (1964). Investment in human capital. *The American Economic Review, 51*, 1–17.

Silva, L. P., Montenegro, A. B., Dias, L. C. F., & Castro, M. A. R. (2015). Educação superior, mobilidade social e expectativa profissional: uma análise à luz da sociologia da educação [Higher education, social mobility and professional expectations: an analysis based on the sociology of education]. *Gestão & Planejamento, 16*, 58–75.

Teichler, U. (2017). Higher education system differentiation, horizontal and vertical. In J. Shin & P. Texeira (Eds.), *Encyclopedia of International Higher Education Systems and Institutions* (pp. 772–778). Springer.

Thompson, J. (2019). Mobility in the middle: bachelor's degree selectivity and the intergenerational association in status in the United States. *Research in Social Stratification and Mobility, 60*, 16–28.

Tomaszewski, W., Perales, F., Xiang, N., & Kubler, M. (2021). Beyond graduation: socio-economic background and post-university outcomes of Australian graduates. *Research in Higher Education, 62*, 26–44.

Trammell, J. (2019). Learning Alone and Academic Capital: An Old Sociological Idea in a New Educational Application. *Journal of College Academic Support Programs, 2*, 62–63.

Vieira, R. S., & Arends-Kuenning, M. (2019). Affirmative action in Brazilian universities: effects on the enrollment of targeted groups. *Economics of Education Review, 73*, 359–375.

Watson, J. (2013). Profitable portfolios: capital that counts in higher education. British *Journal of Sociology of Education, 34*, 412–430.

Winkler, C., & Sriram, R. (2015). Development of a scale to measure academic capital in high-risk college students. *The Review of Higher Education, 38*, 565–587.

# Designing Game-Based Online Courseware for Workplace Chinese Communication in a University Context

Hon Fong Poon, The University of Hong Kong, Hong Kong

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Despite the availability of AI tools for writing, learning Chinese characters, grammar, and writing remains essential for students. It provides a foundational understanding of the language's structure, fostering creativity and allowing for unique self-expression, which are valuable for students' academic achievement and career success, which AI might not fully capture. Self-paced learning is a crucial approach for fostering intrinsic motivation in students, especially in Chinese language learning, which requires extended practice to achieve proficiency. This paper attempts to present the detailed structure and design of an online self-learning courseware, "Game-based Chinese Communication for the Workplace", in the hope that it may provide useful perspectives for international educational academia interested in game-based self-learning resources. Developed with a game-like interface, the courseware features level-based challenges, instant feedback, and reward mechanisms to enhance student motivation and engagement. Set within a diverse company, students (players) assume the role of new employees tasked with solving work-related challenges assigned by superiors and colleagues, all centered on Chinese language application. The question bank comprises five categories, corresponding to five levels, each simulating a different company department through mini-games. The courseware is being developed using Unity as the game development engine, allowing for a dynamic and interactive learning experience. It will be seamlessly integrated into the university's official online learning platform. Expected to launch in September 2026, the courseware will be accessible to all University of Hong Kong students via their student credentials as a voluntary online self-learning resource.

*Keywords:* self-paced learning, online learning, language learning, gamification, student engagement, motivation

iafor
The International Academic Forum
www.iafor.org

## Introduction

To effectively leverage Chinese in academic or workplace settings for communication, strong Chinese writing skills are indispensable. Despite years of study, even native Chinese-speaking students often fall short of the expected proficiency in Chinese writing, particularly in the use of Chinese characters, grammar, and different genres of Chinese writing. For example, each year, following the Hong Kong Diploma of Secondary Education Examination (HKDSE), the Hong Kong Examinations and Assessment Authority (HKEAA) publishes a report detailing the common Chinese character mistakes made by candidates in their responses. These reports consistently reveal that most errors involve frequently used Chinese characters in everyday life, rather than obscure ones. This highlights the importance of grasping the cultural and semantic nuances embedded in Chinese characters, as such knowledge is essential for accurate and effective language use.

Furthermore, a solid grasp of Chinese grammar allows for clear and coherent expression, reduces misunderstandings and aids in the comprehension of complex texts. It also offers linguistic flexibility, enabling individuals to adapt their communication style to different contexts and audiences. These skills are crucial for critical analysis and cultural appreciation in today's interconnected world. As graduates enter the workforce, they often find themselves needing to compose a variety of Chinese documents, ranging from simple emails to detailed report. In crafting these communications, precision in word choice, sentence fluency and adherence to etiquette are crucial.

Despite the availability of AI tools for writing, learning Chinese characters, grammar, and writing remains essential for students. It provides a foundational understanding of the language's structure, fostering creativity and allowing for unique self-expression, which AI might not fully capture. In professional contexts, workplace communication in Chinese demands subtle register choices, socio-pragmatic awareness, and an understanding of institutional dynamics—subtleties that generic AI outputs often miss.

To bridge this gap, this paper introduces "Game-based Chinese Communication for the Workplace", an online self-access courseware designed for students at the University of Hong Kong (HKU). By blending game-based learning (GBL) with authentic workplace scenarios, the courseware provides a voluntary, interactive environment where students gradually master different Chinese language knowledge and build communicative confidence through gamified task-based learning.

## Game-Based Self-Paced Online Courseware

Language learning requires time. Self-paced learning is a crucial approach for fostering intrinsic motivation in students, especially in Chinese language learning, which requires extended practice to achieve proficiency. While self-learning resources are vital for language acquisition outside the classroom, motivating students to engage with online self-learning resources can be challenging.

Gamification is using game-based mechanics, aesthetics and game thinking to engage people, motivate action, promote learning, and solve problems (Kapp, 2012, p. 10). What distinguishes gamification most distinctly from more traditional approaches is the explicit use of competition as a motivational tool. This competitive element is a source of motivation (Nicholson, 2012). The development of the self-paced online courseware "Winning at Work: Game-based Chinese

Communication for the Workplace" is aimed at equipping university students with the Chinese language competencies required to navigate their future professional environments. This online courseware approach enables independent, enjoyable and effective study outside the classroom. By blending technology with game-based activities, the initiative not only improves Chinese proficiency and communication skills but also promotes cognitive development and independent learning. Students can have repeated rounds of training through their personal learning progress, which strengthening their internal motivation rather than pressure on learning Chinese.

## Courseware Structure Design and Game Mechanics

To engage students accustomed to online gaming, this web-based instructional tool adopts a stage-clearing structure, a format familiar to digital gamers, to foster initial interest. As players answer questions to reach the required score threshold for progression, the platform provides immediate corrective feedback and award points upon task completion. This mechanism serves to encourage sustained persistence through higher-difficulty levels or prompt iterative attempts when performance falls short of the passing criteria. The gamified engagement mechanics used in this online courseware include 3 elements: level challenges, instant feedback and reward loops.

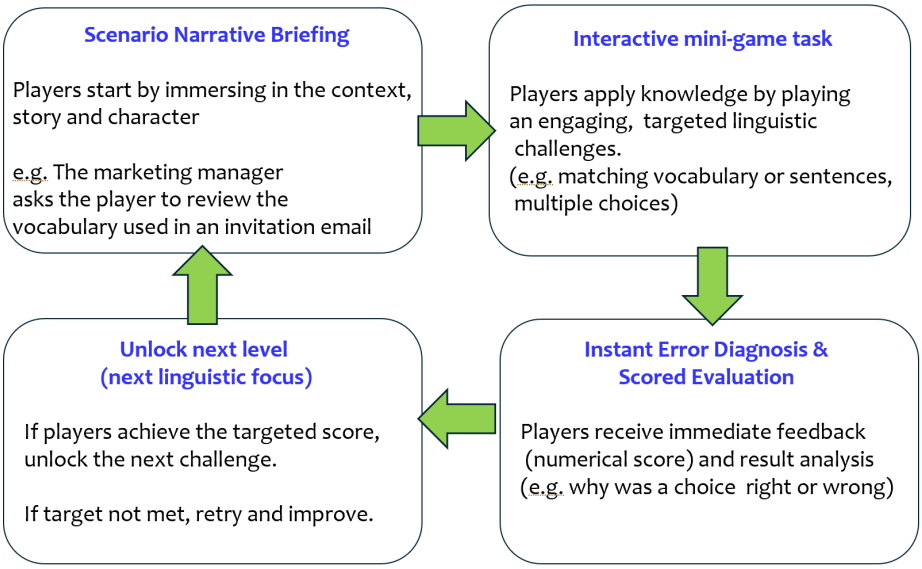
### *Narrative Context and Player Role*

The game story is set in a commercial company, where students will take on the role of newly hired management trainees. Throughout the process, players will handle daily administrative tasks, draft internal documents, negotiate with clients, and use appropriate Chinese language to resolve workplace conflicts, thereby progressing through the levels.

### *Level Challenges*

Since the game setting is a commercial company, each level is structured around an iterative pedagogical game loop:

**Figure 1**
*Game Loop*



### *Five-Level of Workplace Mastery*

The curriculum is organized into five progressive levels. Each level tests specific linguistic skills and pragmatic functions.

**Table 1**
*Five Progressive Levels of Linguistic Focus*

| Level | Core Linguistic Focus | Target Workplace Challenge Focus |
|---|---|---|
| Level 1 | Chinese character | To identify incorrect Chinese characters |
| Level 2 | Vocabulary | To identify correct vocabulary usage |
| Level 3 | Grammar | To determine whether the sentence is grammatically correct or incorrect |
| Level 4 | Etiquette | To determine the most appropriate wording and tone for workplace communication |
| Level 5 | Practical Writing | Master the writing of various practical writing in workplace, such as emails, proposals and invitation letters etc |

For the scenario narrative briefing, there will be a superior or a senior colleague assign an authentic task, for example, ask the player to draft an invitation email to a VIP client, check and identify the grammatical error in a proposal etc. The player navigates a mini-game combining text completion (interactive challenge), pragmatic choice nodes and error identification.

Centered on workplace Chinese, the online courseware presents its five levels in a scaffolded progression of linguistic complexity. The curriculum advances systematically: beginning with foundational skills such as identifying typographical errors in business emails and official documents, moving through the selection of appropriate register-specific vocabulary and idioms, addressing sentence-level syntax, and culminating in advanced capabilities such as evaluating business writing structures and tailoring tone to diverse workplace audiences. This scaffolded approach facilitates incremental skill acquisition, allowing learners to progressively internalize and refine their pragmatic language proficiency.

### *Instant Feedback*

Immediate instructional feedback plays a critical role in facilitating linguistic mastery within gamified environments. At the conclusion of each level, players are provided with real-time performance evaluations along with formative explanations, which assist in diagnosing conceptual gaps and misconceptions in erroneous answers, enabling them to understand the underlying rationale behind incorrect responses. Furthermore, a performance dashboard is displayed at the end of each stage, summarizing the completion rate, accuracy level and completion time. Given the time-constrained design of each level, the task completion rate alongside time-on-task serves as a functional proxy for measuring the learner's automaticity and mastery of the target linguistic focus.

### *Reward Loops*

Rewards serve as a metric to incentivize players to clear levels or re-engage with challenges. Performance is rated on a 3-star scale based on completion rate, accuracy and response speed. Earning at least 2 stars unlocks the next level. Each stage incorporates a dedicated item bank designed to test stage-specific linguistic competencies, with questions generated randomly during each session. For instance, Stage 2 focuses on vocabulary, while Stage 3 targets syntax. Consequently, even if a student successfully clears a level on their first attempt, re-engaging with the platform will present a different subset of questions. This dynamic retrieval mechanism encourages repeated usage and long-term engagement. Furthermore, the item bank can be continuously updated, ensuring the courseware functions as a sustainable, reusable learning tool rather than a single-use educational game.

### Engine Infrastructure and Technical Integration

The courseware is developed using the Unity engine, ensuring cross-platform rendering across desktop web browsers. Also, Unity as the core game logic engine yields high interactivity and provides a rich environment for building complex mini-games natively embedded inside a web frame. The courseware is fully integrated into the university's official online learning platform (HKU OpenEdx platform). All HKU students can use HKU single sign-on (SSO) to get access to this self-paced courseware. The development of the online courseware is still in progress and will be launched in September 2026.

Analytics tracking is also designed, including players learning data, including item-level response time, attempt frequencies, common error patterns and completion rates, which can provide valuable information for continuous courseware refinement and pedagogical research.

## Conclusion

The rapid proliferation of artificial intelligence (AI) and natural language processing (NLP) platforms has ignited a paradigm shift in language education. While AI tools can generate business correspondence on demand, effective workplace communication requires human discretion to evaluate nuance, cultural appropriateness, and emotional resonance. By placing learners in simulated decision-making roles, the courseware cultivates active linguistic judgment rather than passive reliance on automated generation.

Language anxiety often hinders students from writing in formal settings. The gamified corporate environment creates a safe "sandbox" where mistakes carry no real-world professional consequences. Transferring workflow challenges into mini-game environments successfully lowers initial barriers. This encourages trial, reflection and confidence-building. As a voluntary, self-paced tool, the system extends classroom instruction without increasing teaching staff workload. Automated scoring and feedback deliver personal evaluation at scale for large university cohorts. Lastly, building the game-based courseware by using a popular game engine alongside the university system sets a strong example of interactive self-paced learning.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

In the writing process, AI tools Gemini 3.6 Flash and Grammarly are used to help polish the sentences.

## References

Buckley, P., & Doyle, E. (2016). Gamification and student motivation. *Interactive Learning Environments, 24*(6), 1162–1175.

Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education*. Pfeiffer.

Kim, B. (2012). Harnessing the power of game dynamics why, how to, and how not to gamify the library experience. *College & Research Libraries News, 73*(8), 465–469.

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. Penguin.com.

Nicholson, S. (2012, June). *A user-centered theoretical framework for meaningful gamification*. Paper presented at Games+Learning+Society 8.0, Madison, WI.

# Individual Conceptions of Teaching of Prospective Pre-primary Teachers in Relation to the Integration of Educational Content

Ivana Rochovská, Matej Bel University, Slovakia

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

An individual conception of teaching constitutes a system of pedagogical beliefs that shapes teachers' instructional decision-making. This study analyses the individual conceptions of teaching held by prospective pre-primary teachers (N = 66) and examines their relationship to the integration of educational content in instructional planning. The research was conducted in Slovakia, a country with a post-socialist educational context whose education system has been undergoing a long-term transformation towards child-centred approaches. Data were collected using the questionnaire My Individual Conception of Teaching, administered in a pre-test–post-test design during a semester of general didactic training. In addition, students' thematic teaching projects, designed as one-day instructional plans for kindergarten, were analysed through qualitative content analysis. The analytical framework developed in the first part of the study served as the basis for evaluating these projects. Descriptive analysis indicated that students' conceptions of teaching were most strongly oriented towards pragmatist (M_pre = 54.0; M_post = 53.5) and personalist approaches (M_pre = 41.7; M_post = 44.7). A paired samples t-test comparing pre-test and post-test results revealed a statistically significant change only in the personalist conception of teaching ($t(65) = -3.57$, $p < .001$, Cohen's $d = 0.44$). A relationship between individual conceptions of teaching and the degree of educational content integration in thematic projects was identified only in selected aspects of the pragmatist approach, particularly in conceptions of the child's role in education, the organisation of instruction, and assessment. The findings suggest that prospective teachers' pedagogical beliefs may be associated with how integrated learning is planned in pre-primary education.

*Keywords:* individual conceptions of teaching, prospective pre-primary teachers, integration of educational content, content analysis

iafor
The International Academic Forum
www.iafor.org

## Introduction

Students preparing for a career in pre-primary education possess a set of beliefs, values, and expectations concerning the teaching and learning of children. In professional practice, their influence on children's development becomes significant. Even during teaching practice, students may view the child either as an object of their educational influence or, conversely, as an autonomous subject engaged in self-development. These may be understood as two contrasting perspectives that frame what is often referred to as the student's conception of teaching – the underlying reasoning through which they justify their actions in the classroom.

Pre-primary education in the context of the research means the last year of kindergarten before entrance to school, which is compulsory in Slovakia. The Slovak curriculum for early childhood education is a document based on constructivist principles of education. In this framework, the child is viewed as an active, self-developing individual who, however, actualises themselves to the extent and by the means permitted within their culture.

A distinctive feature of the Slovak curriculum for early childhood education is the division of educational content into learning areas (Language and Communication, Mathematics and Information Processing, Human and Nature, Human and Society, Human and the World of Work, Arts and Culture, Health and Physical Activity). This may lead teachers to design educational activities focused on a single learning area. However, in line with current educational requirements, it is important to take an integrated approach to educational content.

## Literature Review

Research into teachers' subjective conceptions has a long tradition. This phenomenon was first examined by Janowski (1980), who described it as a constellation of knowledge, attitudes towards subject matter and textbooks, perceptions of learners, aspirations, expectations, and self-conceptions. The exploration of teachers' thought processes, conceptions, and interpretations was relatively novel, as educational research had until then concentrated primarily on observing and describing in detail the activities of teachers and learners during lessons. Such research, however, was unable to explain why teachers employed certain strategies rather than others (Gavora, 1990). Nor could it clarify why there was no straightforward relationship between the observable activities of teachers and learners and the learning outcomes achieved. This contributed to the recognition of an additional factor influencing teaching effectiveness – namely, the teacher's *individual conception of teaching* (Calderhead, 1987, 1988; Gavora, 1990; Shavelson & Stern, 1981). According to Mareš (2013), the teacher's conception of teaching functions as a *filter*: from official regulations and directives, teachers extract those elements they accept and further adapt them. Through this *filter*, teachers also evaluate their own pedagogical activity, the activities and outcomes of learners, as well as the practices of colleagues, parents, and even the wider public.

An individual conception of teaching in a student may be expressed as an inclination towards the perspectives of existing pedagogical conceptions, particularly in the areas of objectives, content, methods, forms, the roles of participants in teaching, and the assessment of learning outcomes (Rochovská, 2026; Rochovská & Brozmanová, 2025). The study operates with the perspectives of three pedagogical approaches: conservatism, pragmatism, and personalism.

*Conservatism* conceptualises education as the transmission of cultural heritage, emphasising the formative role of the social environment, the authority of the teacher, and the development

of discipline and perseverance in the child. *Pragmatism* understands education as a process of learning through interaction and experience, focusing on the acquisition of practical competencies, the integration of learning with real-life contexts, and the development of learners' abilities to construct and apply knowledge. *Personalism* views education as the facilitation of individual development, prioritising the learner's autonomy, personal growth, and self-realisation, with the teacher acting as a guide through non-directive pedagogical approaches (Rochovská, 2026).

According to Kerry (2015), integrated teaching enables content to be linked through the exploration of clear and relevant connections across the curriculum. It encourages the use and application of what has been taught and learned in new and different ways and provides opportunities to deepen learning, for example, through answering big questions, exploring a problem, solving a problem, or completing a final project. Learning beyond subject boundaries provides opportunities for students to experience deep, challenging and relevant learning.

Rakoušová (2008) defines integration as a process in which the content of subjects derived from the sciences is interconnected and combined. This content is designed to create new, functional educational content characterised by stronger interconnections. The goal of integrated educational content is to pursue the objectives of all educational areas involved in this process.

According to the definition provided by Beane (1997), curriculum integration encompasses four main aspects: integration of experiences, social interaction, interaction of knowledge, and integration as curriculum design. By analysing each of these aspects, one can identify the interconnections that together form a comprehensive theory of curricular integration. The author further notes that in curriculum integration, thematic areas are grounded in everyday experiences and real life, thereby enabling young people to critically examine current social issues and engage in social activities according to their needs. This approach aims to make learning more meaningful by focusing on real-world topics that stimulate critical thinking and social engagement (Beane, 1997).

## Research Problem and Research Questions

In the context of this research, the integration of educational content refers to the interconnection of various subject areas into a coherent whole, where children learn not only specific facts but also the relationships between them, which helps them understand complex issues and develop holistic thinking. Based on the ideas outlined above, there arose a need to examine how the individual conceptions of teaching held by students—future early childhood educators—relate to their ability to connect and integrate educational content.

The following research questions emerged from the research problem: RQ1: Which aspects of pedagogical conceptions are reflected in the individual conceptions of teaching held by students of pre-primary education? RQ2: Is there a difference between students' conceptions of teaching at the beginning and at the end of the semester? RQ3: Is there a relationship between individual conceptions of teaching (i.e. the inclination towards particular aspects of pedagogical conceptions) and the implementation of integrated educational content in the teaching projects they developed?

# Methodology

## Research Design

To achieve the research objective—analysing the individual conceptions of teaching held by prospective pre-primary teachers and examining their relationship to the integration of educational content in instructional planning—a combined quantitative and qualitative research design was employed.

## Research Sample

The research sample consisted of 66 prospective pre-primary teachers (convenience sample) in the second year of a bachelor's degree programme at the Faculty of Education. The sample also included 66 education projects that they developed during the course Theory of Teaching in Pre-primary Education.

## Research Methods and Procedures

The individual conceptions of teaching held by students were assessed using a questionnaire. The foundations for the development of the questionnaire have been published in the author's previous works (Rochovská, 2026; Rochovská & Brozmanová, 2025). Each respondent was assigned a score consisting of a combination of values reflecting their inclination towards the conceptions of conservatism, pragmatism, and personalism.

The questionnaire (Rochovská, 2025), developed in previous research, consisted of 30 items using a Likert scale. For each statement—alternately corresponding to the perspectives of conservatism, personalism, and pragmatism (relating to objectives, content, teaching strategies and methods, the roles of the teacher and the child, the organisation of teaching, and assessment)—respondents indicated their level of agreement on a scale. Points were assigned according to their responses. The scale was as follows: strongly agree (4 points), mostly agree (3 points), mostly disagree (2 points), strongly disagree (1 point), and unable to express an opinion (0 points).

The questionnaire also included an additional 7 items with a single-response format. These addressed the same areas as the Likert-scale items; however, respondents were required to choose one of three options corresponding to the perspectives of conservatism, pragmatism, or personalism. Each selected option was awarded 4 points.

The questionnaire was administered to students at the beginning of the semester (pre-test) and at the end of the semester (post-test).

The internal consistency of the scales was assessed using Cronbach's alpha. The conservatism dimension reached a value of $\alpha = 0.650$, the pragmatism dimension $\alpha = 0.501$, and the personalism dimension $\alpha = 0.595$. These values are not high, indicating limited internal consistency of the scales. Therefore, the subscales should not be interpreted as internally consistent reflective measures, but rather as theoretically grounded groupings of indicators capturing broader pedagogical orientations. The subscales were therefore interpreted as composite indicators of broader pedagogical orientations rather than internally consistent scales.

The implementation of integrated educational content in the developed teaching projects was examined through an analysis of these projects based on defined criteria. The analytical framework developed in the first part of the study served as the basis for evaluating these projects. A manual for coding students' work with a focus on the integration of educational content was developed.

During the development of the manual, it became evident that content-integrated teaching also requires a shift in the approach to the child as a subject of learning. This was incorporated into the categories relating to the role of the teacher and the child, as well as teaching methods and forms, the organisation of teaching, and the assessment of learning outcomes. The manual is presented in Table 1.

**Table 1**
*A Manual for Coding Students' Work With a Focus on the Integration of Educational Content*

| Category | Criterion | Score |
|---|---|---|
| 1. Integrated Objectives | Objectives integrated from all areas of pre-primary education. | 4 |
| | Objectives integrated from six to five areas of pre-primary education. | 3 |
| | Objectives integrated from four to three areas of pre-primary education. | 2 |
| | Objectives integrated from two areas of pre-primary education. | 1 |
| | Objectives from one area of pre-primary education. | 0 |
| 2. Integrated Content | Content integrated from all areas of pre-primary education. | 4 |
| | Content integrated from six to five areas of pre-primary education. | 3 |
| | Content integrated from four to three areas of pre-primary education. | 2 |
| | Content integrated from two areas of pre-primary education. | 1 |
| | Content from one area of pre-primary education. | 0 |
| 3. Methods and Forms | Activating methods used in all activities. | 4 |
| | Predominantly activating methods in most activities. | 3 |
| | Balanced application of activating and transmissive activities. | 2 |
| | Transmissive methods in almost all activities. | 1 |
| | Transmissive methods in all activities. | 0 |
| 4. Teacher's Role | The teacher acts as a facilitator in all activities. | 4 |
| | The teacher is more of a facilitator than an authority. | 3 |
| | The teacher is predominantly an authority but allows children to co-decide in at least two activities. | 2 |
| | The teacher is predominantly an authority but allows children to co-decide in at least one activity. | 1 |
| | The teacher is an authority, assigns tasks and controls. | 0 |
| 5. Child's Role | The child is an active participant in learning. | 4 |
| | The child is predominantly an active participant in learning. | 3 |
| | In several activities, the child is more active than passive. | 2 |
| | The child is an object of instruction in almost all activities. | 1 |
| | The child is an object of instruction in all activities. | 0 |

| | | |
|---|---|---|
| 6. Organisation of Teaching | Individualised tasks and group work in all activities. | 4 |
| | Individualised tasks and group work in almost all activities. | 3 |
| | Equal representation of whole-class, individual, and group work. | 2 |
| | Whole-class teaching in almost all activities. | 1 |
| | Whole-class teaching in all activities. | 0 |
| 7. Assessment | Formative assessment focused on the child's personal development. | 4 |
| | Predominantly formative assessment focused on the child's personal development. | 3 |
| | Combination of summative and formative assessment. | 2 |
| | Predominantly summative assessment. | 1 |
| | Summative assessment. | 0 |

Deductive coding, based on the predefined manual, was supported by artificial intelligence (ChatGPT.org tool). For each analysed project, the prompt included a request for the project to be coded according to the manual (Table 1). To enhance reliability, all AI-generated coding outputs were systematically reviewed and, where necessary, corrected by the author. This procedure ensured consistency with the coding manual and reduced potential bias associated with automated analysis.

### Data Analysis

Descriptive statistics were used to evaluate respondents' inclinations towards the perspectives of individual pedagogical conceptions. Differences between the pre-test and post-test were analysed using paired samples t-tests (given the approximately interval nature of aggregated scale scores) and complemented by non-parametric Wilcoxon signed-rank tests at the item level. Correlations between individual conceptions of teaching and the implementation of integrated educational content were analysed using Spearman's rho. Given the ordinal nature of Likert-scale data and the low internal consistency of the scales, non-parametric Wilcoxon signed-rank tests were additionally employed at the item level.

## Results

### Results of the Questionnaire

Table 2 presents the results of the descriptive statistics for respondents' inclinations towards the perspectives of pedagogical conceptions in the pre-test and post-test.

**Table 2**
*The Respondents' Inclinations Towards the Perspectives of Pedagogical Conceptions in the Pre-test and Post-test*

| | N | Mean | Median | SD |
|---|---|---|---|---|
| Perspectives of conservatism (pre-test) | 66 | 30.3 | 30.0 | 4.42 |
| Perspectives of conservatism (post-test) | 66 | 31.3 | 31.0 | 5.47 |
| Perspectives of pragmatism (pre-test) | 66 | 54.0 | 55.0 | 6.44 |
| Perspectives of pragmatism (post-test) | 66 | 53.5 | 53.0 | 6.18 |
| Perspectives of personalism (pre-test) | 66 | 41.7 | 41.0 | 6.98 |
| Perspectives of personalism (post-test) | 66 | 44.7 | 44.0 | 6.87 |

The results indicate that respondents showed the strongest inclination towards pragmatist and personalist conceptions of teaching, both in the pre-test and post-test. Pragmatism achieved the highest mean scores (M_pre = 54.0; M_post = 53.5), suggesting a stable preference for approaches emphasising practical relevance and real-life application.

Personalism showed a noticeable increase (M_pre = 41.7; M_post = 44.7), indicating a shift towards more child-centred perspectives. In contrast, conservatism remained the least preferred orientation, although a slight increase was observed (M_pre = 30.3; M_post = 31.3).

**Table 3**
*Comparison of Respondents' Answers in the Pre-test and Post-test Using a Paired Samples t-Test*

| | | | **statistic** | **df** | **p** | | **Effect Size** |
|---|---|---|---|---|---|---|---|
| Perspectives of conservatism (pre-test) | Perspectives of conservatism (post-test) | Student's t | -1.604 | 65.0 | 0.114 | Cohen's d | -0.1974 |
| Perspectives of pragmatism (pre-test) | Perspectives of pragmatism (post-test) | Student's t | 0.590 | 65.0 | 0.557 | Cohen's d | 0.0727 |
| Perspectives of personalism (pre-test) | Perspectives of personalism (post-test) | Student's t | -3.574 | 65.0 | < .001 | Cohen's d | -0.4399 |

The paired samples t-test revealed a statistically significant difference only for personalism ($t(65) = -3.57$, $p < .001$), with a small-to-moderate effect size ($d = 0.44$). This indicates that the course contributed to a measurable shift towards more child-centred conceptions of teaching. No statistically significant changes were found for conservatism or pragmatism, suggesting that these orientations remained relatively stable over the course of the semester.

Table 4 presents the correlations between inclinations towards the perspectives of pedagogical conceptions in the pre-test and post-test (Pearson correlation analysis).

**Table 4**
*The Correlations Between Inclinations Towards the Perspectives of Pedagogical Conceptions in the Pre-test and Post-test (Pearson Correlation Analysis)*

| | | Conservatism (pretest) | Pragmatism (pretest) | Personalism (pretest) |
|---|---|---|---|---|
| Pragmatism (pretest) | Pearson's r | 0.264 | — | |
| | df | 64 | — | |
| | p-value | 0.032 | — | |
| Personalism (pretest) | Pearson's r | -0.303 | -0.850 | — |
| | df | 64 | 64 | — |
| | p-value | 0.013 | < .001 | — |

The correlation analysis shows several statistically significant relationships between the examined conceptions. Notably, negative correlations were identified between personalism and both conservatism and pragmatism, indicating that these perspectives may represent partially contrasting orientations.

The strong negative correlation between pragmatism and personalism suggests that, within this sample, these orientations function as distinct and potentially competing frameworks rather than complementary approaches.

Due to the low internal consistency of the scales, item-level analysis was conducted (Table 5). Statistically significant values are indicated in bold.

**Table 5**

*Comparison of Pre-test and Post-test Results for Individual Items (Paired Samples t-Test)*

| | | | | |
|---|---|---|---|---|
| 1. At school, we help children learn about our traditions and what is right and wrong. (pre-test) | 1. At school, we help children learn about our traditions and what is right and wrong. (post-test) | Wilcoxon W | 187.5 [a] | 0.459 |
| 2. Learning should be connected to everyday life so children can understand it better. (pre-test) | 2. Learning should be connected to everyday life so children can understand it better. (post-test) | Wilcoxon W | 45.5 [b] | 0.594 |
| 3. At school, we should help each child grow and develop in their own way. (pre-test) | 3. At school, we should help each child grow and develop in their own way. (post-test) | Wilcoxon W | 85.0 [d] | 0.143 |
| **4. It is important that children learn to respect rules and adults at school. (pre-test)** | **4. It is important that children learn to respect rules and adults at school. (post-test)** | **Wilcoxon W** | **352.5 [e]** | **0.009** |
| 5. School should help children learn how to solve real-life problems. (pre-test) | 5. School should help children learn how to solve real-life problems. (post-test) | Wilcoxon W | 13.5 [f] | 0.530 |
| 6. We should take into account each child's needs and interests. (pre-test) | 6. We should take into account each child's needs and interests. (post-test) | Wilcoxon W | 93.5 [g] | 0.385 |
| 7. At school, we help children keep good habits and value traditions. (pre-test) | 7. At school, we help children keep good habits and value traditions. (post-test) | Wilcoxon W | 149.5 [h] | 0.708 |
| **8. Children should learn how to learn and how to gain new knowledge. (pre-test)** | **8. Children should learn how to learn and how to gain new knowledge. (post-test)** | **Wilcoxon W** | **396.0 [i]** | **0.025** |
| **9. An important goal of education is for each child to develop their own potential. (pre-test)** | **9. An important goal of education is for each child to develop their own potential. (post-test)** | **Wilcoxon W** | **55.0 [j]** | **0.017** |
| **10. Teaching should include activities that help children become disciplined and persistent. (pre-test)** | **10. Teaching should include activities that help children become disciplined and persistent. (post-test)** | **Wilcoxon W** | **402.0 [k]** | **0.049** |
| 11. Learning should mainly be based on children's own activities, exploring, discovering, and experimenting. (pre-test) | 11. Learning should mainly be based on children's own activities, exploring, discovering, and experimenting. (post-test) | Wilcoxon W | 32.0 [l] | 0.076 |
| 12. In school, children should be given the chance to make their own decisions and take responsibility for their learning. (pre-test) | 12. In school, children should be given the chance to make their own decisions and take responsibility for their learning. (post-test) | Wilcoxon W | 288.0 [m] | 0.628 |
| 13. Children should learn to concentrate and work carefully according to given rules. (pre-test) | 13. Children should learn to concentrate and work carefully according to given rules. (post-test) | Wilcoxon W | 418.0 [n] | 0.156 |
| 14. It is important that children learn how to find and use information. (pre-test) | 14. It is important that children learn how to find and use information. (post-test) | Wilcoxon W | 141.5 [d] | 0.610 |
| 15. Teaching should be flexible and adapted to each child's needs. (pre-test) | 15. Teaching should be flexible and adapted to each child's needs. (post-test) | Wilcoxon W | 162.5 [o] | 0.695 |

| | | | | |
|---|---|---|---|---|
| 16. The teacher's role is to pass on knowledge to children. (pre-test) | 16. The teacher's role is to pass on knowledge to children. (post-test) | Wilcoxon W | 297.0 [c] | 0.148 |
| 17. Children learn best when they are active and involved. (pre-test) | 17. Children learn best when they are active and involved. (post-test) | Wilcoxon W | 14.0 [p] | 0.484 |
| 18. The teacher should guide children on their learning journey. (pre-test) | 18. The teacher should guide children on their learning journey. (post-test) | Wilcoxon W | 56.0 [r] | 0.835 |
| 19. The teacher's role is mainly to help children settle into the school environment. (pre-test) | 19. The teacher's role is mainly to help children settle into the school environment. (post-test) | Wilcoxon W | 145.5 [s] | 0.100 |
| 20. Children learn best when they can try things themselves and learn from mistakes. (pre-test) | 20. Children learn best when they can try things themselves and learn from mistakes. (post-test) | Wilcoxon W | 36.0 [t] | 0.802 |
| 21. Every child should have the chance to work at their own pace. (pre-test) | 21. Every child should have the chance to work at their own pace. (post-test) | Wilcoxon W | 84.0 [d] | 0.133 |
| 22. The teacher's role is to guide children using proven methods. (pre-test) | 22. The teacher's role is to guide children using proven methods. (post-test) | Wilcoxon W | 253.0 [s] | 0.424 |
| **23. Children should learn new things that build on what they already know. (pre-test)** | **23. Children should learn new things that build on what they already know. (post-test)** | **Wilcoxon W** | **143.0 [n]** | **0.002** |
| 24. The teacher should encourage children to set their own learning goals. (pre-test) | 24. The teacher should encourage children to set their own learning goals. (post-test) | Wilcoxon W | 246.0 [m] | 0.728 |
| 25. All children should try to keep up with the pace set by the teacher. (pre-test) | 25. All children should try to keep up with the pace set by the teacher. (post-test) | Wilcoxon W | 399.0 [n] | 0.271 |
| 26. At school, children should be able to work together when solving tasks. (pre-test) | 26. At school, children should be able to work together when solving tasks. (post-test) | Wilcoxon W | 92.0 [d] | 0.207 |
| 27. Each child should have their own learning plan and work at their own pace. (pre-test) | 27. Each child should have their own learning plan and work at their own pace. (post-test) | Wilcoxon W | 285.0 [u] | 0.614 |
| 28. The teacher should assess each child's knowledge fairly. (pre-test) | 28. The teacher should assess each child's knowledge fairly. (post-test) | Wilcoxon W | 170.0 [v] | 0.631 |
| 29. When assessing, the teacher should focus on children's work, for example their projects. (pre-test) | 29. When assessing, the teacher should focus on children's work, for example their projects. (post-test) | Wilcoxon W | 246.5 [k] | 0.366 |
| 30. Assessment should focus on each child's progress and how much they have improved. (pre-test) | 30. Assessment should focus on each child's progress and how much they have improved. (post-test) | Wilcoxon W | 56.0 [r] | 0.428 |

Given the ordinal nature of Likert-scale data and the low internal consistency of the scales, non-parametric Wilcoxon signed-rank tests were additionally employed at the item level. Statistically significant differences were found only in selected items, particularly those related to: respect for rules and authority (Item 4), learning how to learn (Item 8), development of the child's potential (Item 9), discipline and persistence (Item 10), building on prior knowledge (Item 23).

These results suggest that changes occurred at the level of specific beliefs rather than across entire conceptual frameworks, supporting the earlier finding of partial rather than systemic change.

## Results of the Content Analysis

The results of the content analysis of the students' projects are presented in Table 6.

**Table 6**
*Correlation Between Individual Conceptions of Teaching, As Measured by the Questionnaire, and the Teaching Projects Developed by the Students (Spear-Man's Rho)*

| | | 1. Integrated Objectives | 2. Integrated Content | 3. Methods and Forms | 4. Teacher' s Role | 5. Child' s Role | 6. Organisation of Teaching | 7. Assessment |
|---|---|---|---|---|---|---|---|---|
| Conservatism (pretest) | Correlation Coefficient | ,074 | ,049 | ,180 | ,049 | ,024 | ,089 | ,116 |
| | Sig. (2-tailed) | ,556 | ,694 | ,148 | ,694 | ,846 | ,475 | ,352 |
| Pragmatism (pretest) | Correlation Coefficient | -,131 | -,133 | -,156 | -,133 | -,181 | -,184 | -,172 |
| | Sig. (2-tailed) | ,294 | ,286 | ,211 | ,286 | ,145 | ,138 | ,167 |
| Personalism (pretest) | Correlation Coefficient | ,101 | ,114 | ,104 | ,114 | ,096 | ,075 | ,061 |
| | Sig. (2-tailed) | ,418 | ,364 | ,405 | ,364 | ,441 | ,547 | ,625 |
| Conservatism (posttest) | Correlation Coefficient | -,173 | -,174 | -,107 | -,174 | -,230 | -,181 | -,123 |
| | Sig. (2-tailed) | ,164 | ,161 | ,391 | ,161 | ,063 | ,147 | ,327 |
| Pragmatism (posttest) | Correlation Coefficient | -,134 | -,132 | -,062 | -,132 | **-,328** | **-,321** | **-,289** |
| | Sig. (2-tailed) | ,283 | ,289 | ,619 | ,289 | **,007** | **,009** | **,019** |
| Personalism (posttest) | Correlation Coefficient | ,068 | ,042 | ,037 | ,042 | ,143 | ,035 | ,158 |
| | Sig. (2-tailed) | ,588 | ,738 | ,768 | ,738 | ,253 | ,782 | ,205 |

The correlation analysis between questionnaire results and teaching projects revealed mostly weak and non-significant relationships. However, statistically significant correlations were identified for pragmatism in the post-test, specifically in: child's role ($\rho = -.328$, $p = .007$), the organisation of teaching ($\rho = -.321$, $p \approx .012$), and assessment ($\rho = -.289$, $p = .019$).

These negative correlations indicate that higher scores on pragmatist orientations were associated with lower levels of integration in these aspects of instructional design.

No significant relationships were found for conservatism or personalism, indicating that the translation of beliefs into practice may be limited or inconsistent.

## Discussion

The findings indicate that the individual conceptions of teaching held by prospective pre-primary teachers are predominantly oriented towards pragmatism and personalism, which aligns with contemporary child-centred and constructivist approaches in early childhood education. This is consistent with the broader curricular transformation in Slovakia towards learner-centred education.

A statistically significant shift was observed only in the personalist dimension, suggesting that the course may have contributed to strengthening beliefs related to the child as an active subject of learning. However, the absence of significant changes in pragmatism and conservatism

indicates that some pedagogical beliefs may be more stable and resistant to short-term intervention.

The relationship between conceptions of teaching and the implementation of integrated educational content was found to be limited and selective. Significant associations were identified only within the pragmatist dimension in the post-test, particularly in areas related to the child's role, organisation, and assessment. Interestingly, these relationships were negative, which may suggest that a stronger focus on practical problem-solving does not necessarily translate into integrated curriculum design.

These findings point to a gap between beliefs and practice, a phenomenon widely discussed in teacher education research. Although students may endorse progressive pedagogical ideas, their ability to operationalise these ideas in instructional planning appears to be constrained.

This finding is consistent with previous research highlighting the discrepancy between teachers' espoused beliefs and their actual classroom practices (e.g. Calderhead, 1988).

## Conclusion

This study examined the individual conceptions of teaching held by prospective pre-primary teachers and their relationship to the integration of educational content in instructional planning. The findings indicate that students predominantly favour pragmatist and personalist approaches, with a significant shift observed only in the personalist dimension. Overall, however, pedagogical beliefs remained relatively stable over the course of the semester.

The relationship between beliefs and instructional practice was found to be limited and selective, suggesting that the endorsement of particular pedagogical conceptions does not necessarily translate into their application in teaching design. This highlights the need for teacher education programmes to strengthen the connection between theory and practice, particularly through reflective activities that support the alignment of beliefs and instructional decisions.

The study is limited by the low internal consistency of the scales, the relatively small and context-specific sample, and the use of AI-assisted coding. Future research should focus on improving measurement reliability, employing longitudinal designs, and combining self-report data with observational methods to better capture the relationship between beliefs and practice.

## Acknowledgements

This text was created under the national VEGA project No. 1/0486/24 entitled "Research on teachers' potential and analysis of curriculum documents from the aspect of integration of educational contents of primary education", and national KEGA project No. 001UBM-4/2025, "Letter Games in Early Childhood Education: Development of Methodological Materials Integrating Educational Areas".

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The AI-generation technology (ChatGPT.org tool) was used for the deductive analyses of the students' teaching projects according to the manual, and for improving the language and readability of the paper.

## References

Calderhead, J. (1987). *Exploring teachers' thinking.* Cassell.

Calderhead, J. (1988). *Teachers' thought processes.* The Falmer Press.

Gavora, P. (1990). Učiteľova individuálna koncepcia vyučovania [The teacher's individual teaching concept]. *Pedagogická revue, 42*(3), 209–222.

Janowski, A. (1980). *Social psychology and educational issues*. Ossolineum.

Mareš, J. (2013). *Pedagogická psychologie* [Educational psychology]. Portál.

Rakoušová, A. (2008). *Integrace obsahu vyučování* [Integration of teaching content]. Grada Publishing.

Rochovská, I. (2025). *My Individual Teaching Conception (Questionnaire).* https://docs.google.com/forms/d/1fbgs59XgPdpEpblv5C2VMOzUTqLTjhVYQzrtE0zWLrs/preview

Rochovská, I. (2026). Individual teaching philosophy of parents of homeschooled children in comparison with primary school teachers. *Pedagogika-Pedagogy, 98*(2), 217–228. https://doi.org/10.53656/ped2026-2.04

Rochovská, I., & Brozmanová, M. (2025). Teaching approaches in preschool and elementary school students. In *EDULEARN25 conference proceedings: 17th International Conference on Education and New Learning Technologies* (pp. 2914–2923). IATED. https://doi.org/10.21125/edulearn.2025.0796

Shavelson, R. J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgements, decisions and behaviour. *Review of Educational Research, 51*(4), 455–498. https://doi.org/10.2307/1170362

**Contact email:** ivana.rochovska@umb.sk

# Beyond GenAI Adoption: New Horizons for the Ethos of Teaching in Higher Education

Adma Gama-Krummel, University of Rochester, United States

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

The rapid integration of generative artificial intelligence (GenAI) into higher education has generated extensive research on technological adoption, academic integrity, and pedagogical innovation. Less attention has been given to how GenAI challenges the professional commitments through which faculty interpret responsible teaching. This conceptual literature review examined recurring professional tensions in faculty engagement with GenAI and developed a typology of faculty professional ethos. Following the screening of 357 records, 52 publications available online by December 31, 2025, were retained for conceptual synthesis; 41 provided direct evidence for four tensions: efficiency and scalability versus professional judgment and accountability; innovation and experimentation versus integrity, reliability, and governance; instructional continuity versus pedagogical redesign; and technological augmentation versus the human and relational meaning of teaching. The tensions were organized along two dimensions: generative possibility versus protective responsibility and operational practice versus pedagogical meaning and relationships. Their intersection produced four orientations: the Instrumental Optimizer, Ethical Gatekeeper, Pedagogical Re-Designer, and Critical Humanist. These orientations are overlapping, context-sensitive constructs rather than fixed faculty types or developmental stages. By shifting attention from whether faculty adopt GenAI to how they negotiate responsibility, the typology offers a framework for faculty development, institutional policy, and empirical research on GenAI-mediated teaching.

*Keywords:* generative artificial intelligence, higher education, faculty professional ethos, pedagogy, academic integrity

iafor
The International Academic Forum
www.iafor.org

## Introduction

Generative artificial intelligence (GenAI) supports instructional preparation, feedback, assessment, and administrative work while enabling students to generate text, code, explanations, and alternative solutions. These capabilities invite experimentation but also raise concerns about accuracy, authorship, privacy, academic integrity, assessment validity, and dependence on generated content (Ghimire et al., 2024; Lee et al., 2024; Yusuf et al., 2024).

Research on faculty engagement with educational technology has often used adoption models that emphasize usefulness, ease of use, social influence, and facilitating conditions (Davis, 1989; Venkatesh et al., 2003). These factors remain relevant to GenAI (Mah & Groß, 2024; Miranda & Chamorro-Mera, 2026), but adoption does not fully explain professional judgment. Faculty may welcome GenAI for lesson planning yet reject automated grading, or restrict AI-generated writing while designing activities in which students critique generated responses. Such decisions vary by task, consequence, and educators' interpretations of responsibility (Ellis et al., 2025; Farazouli et al., 2025; Kahn et al., 2025).

This paper examines GenAI as a site where faculty professional ethos becomes visible. Professional ethos refers here to the normative commitments, epistemic assumptions, and pedagogical responsibilities through which faculty interpret responsible teaching. It includes judgments about what may be delegated, what requires human oversight, what students need to learn, and which relationships and forms of judgment should remain central.

This shift in emphasis matters because apparently similar practices can express different professional commitments. A restriction on GenAI may protect assessment validity, student privacy, foundational learning, or the relational meaning of feedback. Likewise, encouraging GenAI may pursue efficiency, disciplinary preparation, critical experimentation, or additional time for student interaction. Treating these decisions simply as adoption or resistance obscures the reasons educators give for them and the responsibilities they are trying to preserve.

The study asks: What recurring professional tensions characterize faculty engagement with GenAI in higher education, and what orientations of professional ethos emerge from these tensions? It contributes a four-part typology—the Instrumental Optimizer, Ethical Gatekeeper, Pedagogical Re-Designer, and Critical Humanist. The typology does not rank or permanently classify educators; it provides a vocabulary for examining how faculty emphasize different responsibilities across tasks and contexts.

## Methodological Approach

This study used a conceptual literature review to identify recurring concepts and relationships across methodologically diverse scholarship (Kennedy, 2007). Integrative review principles supported the synthesis of empirical, conceptual, and policy-oriented publications for conceptual clarification and theory development (Snyder, 2019; Torraco, 2005).

### Search and Selection

Iterative searches were conducted through DiscoverUR, Scopus, ERIC, and Google Scholar. Search combinations joined terms for GenAI (e.g., "generative AI," "ChatGPT," and "large language models") with terms for higher education faculty, teaching, assessment, professional

roles, agency, ethics, integrity, AI literacy, and professional development. Citation tracking of recent reviews and faculty-focused studies supplemented database searches.

The search produced 357 records for title-and-abstract screening. Publications were eligible when they addressed faculty uses or perceptions of GenAI; pedagogical or assessment redesign; professional agency or roles; academic integrity and ethics; AI literacy and development; institutional governance; or relational dimensions of teaching. Empirical studies, conceptual scholarship, reviews, and selected competency or policy frameworks were included when they contributed directly to faculty professional responsibility. The cutoff was December 31, 2025; articles assigned to later issues were included only if formally published online by that date. Full-text assessment retained 52 publications, of which 41 provided the most direct evidence for the tensions in Table 1. The remaining publications supplied theoretical, methodological, contextual, or policy support.

Publications focused exclusively on student adoption, technical system performance, or settings outside higher education were excluded unless they offered a framework necessary for interpreting faculty practice. Because the review sought conceptual development rather than effect estimation, methodological diversity was retained deliberately. The distinction between the 52-publication conceptual corpus and the 41 directly informing publications prevents background frameworks from being presented as equivalent to evidence about faculty practice.

## Conceptual Analysis

Analytic notes recorded how each publication described faculty uses of GenAI, emphasized opportunities or concerns, assigned professional responsibility, and framed the purposes of teaching. Comparison focused on efficiency, assessment, integrity, pedagogical judgment, agency, institutional responsibility, relationships, and faculty roles.

Analysis moved recursively among publications, provisional categories, and the developing framework. Convergent and contradictory claims were compared, and related concerns were grouped rather than treated as predetermined codes (Marshall et al., 2022). Four tensions emerged and were examined along two dimensions: generative possibility versus protective responsibility, and operational practice versus pedagogical meaning and relationships. A typology-oriented design then differentiated recurring configurations within these dimensions (Jaakkola, 2020). The resulting orientations are interpretive, overlapping constructs rather than predictive profiles.

The analysis did not count the number of favorable or unfavorable statements as evidence of prevalence. Instead, it examined how claims about GenAI were connected to conceptions of faculty responsibility. Each directly informing publication was assigned to the tension to which it made its clearest contribution, while analytic notes retained its secondary connections. Provisional orientations were repeatedly compared against studies that combined enthusiasm and caution, thereby testing whether the framework could accommodate overlap rather than reproduce a binary adoption model.

## Conceptual Findings: From Professional Tensions to Ethos Orientations

Faculty engagement in the reviewed literature did not divide neatly into adoption and resistance. Educators often recognized practical or pedagogical possibilities while questioning

reliability, integrity, student learning, or human judgment. The synthesis first identifies four tensions and then shows how their intersection produces four orientations of professional ethos.

## Recurring Professional Tensions

Table 1 places each of the 41 directly informing publications under its primary tension. The tensions are not binary choices; they name commitments that faculty may balance or prioritize differently across contexts.

**Table 1**
*Recurring Professional Tensions Informing the Conceptual Typology of Faculty Engagement With GenAI*

| Professional tension | Connection to faculty professional ethos | Publications primarily informing the tension |
|---|---|---|
| **Efficiency and scalability versus professional judgment and accountability** | Whether automation extends faculty capacity or transfers consequential instructional judgment to GenAI. | Agostini and Picasso (2024); Al-Dosari (2023); Alammari (2024); Cordero et al. (2025); Gaber et al. (2023); Gong and Xiao (2024); Hennekeuser et al. (2024); Miranda and Chamorro-Mera (2026); de Obesso et al. (2025); Pang et al. (2024); Soodan et al. (2024); Wainaina and Sun (2025) |
| **Innovation and experimentation versus integrity, reliability, and governance** | How experimentation is conditioned by authorship, validity, privacy, transparency, bias, accuracy, and policy. | Al-Mughairi and Bhaskar (2023); Alrayes et al. (2024); Deroncele-Acosta et al. (2025); Ghimire et al. (2024); Kiryakova and Angelova (2023); Kumar (2023); Maistry and Singh (2025); Simelane and Kittur (2025); Verano-Tacoronte et al. (2025); Yusuf et al. (2024) |
| **Instructional continuity versus pedagogical redesign** | Whether GenAI supports existing practice or prompts changes in assessment, curriculum, AI literacy, and epistemic agency. | Ali (2024); Barrett and Pack (2023); Cha et al. (2024); Ellis et al. (2025); Kahn et al. (2025); Lau and Guo (2023); Lee et al. (2024); Nikoçeviq-Kurti and Bërdynaj-Syla (2024); Ouyang and Jiao (2021); Rajabi et al. (2023); Salhab (2024); Sjöberg and Bergdahl (2025) |
| **Technological augmentation versus the human and relational meaning of teaching** | Whether GenAI strengthens teaching or weakens agency, dialogue, mentorship, trust, judgment, and disciplinary expertise. | Farazouli et al. (2025); Hasan et al. (2024); Jose and Jose (2024); Mah and Groß (2024); McGrath et al. (2023); Ofosu-Ampong (2024); Pisica et al. (2023) |

*Note.* Publications are placed under their primary contribution, although many inform multiple tensions. Miranda and Chamorro-Mera appears in a 2026 issue but was published online on October 2, 2025.

### *Efficiency and Scalability Versus Professional Judgment and Accountability*

GenAI can accelerate lesson planning, resource creation, feedback, grading, and administration (Agostini & Picasso, 2024; Cordero et al., 2025). It can also make individualized support more feasible in large classes. Yet speed does not establish accuracy, fairness, or pedagogical value. Generated materials still require evaluation against disciplinary standards, learning objectives, and knowledge of particular students. In assessment, the issue is not only whether a system can produce feedback or a score but who remains answerable for the judgment. Efficiency becomes professionally valuable when it extends faculty capacity without obscuring the labor of verification or transferring consequential accountability to a system.

### *Innovation and Experimentation Versus Integrity, Reliability, and Governance*

Faculty experiment with GenAI while confronting authorship, privacy, hallucination, misconduct, and assessment-validity concerns (Ghimire et al., 2024; Simelane & Kittur, 2025). Their responsibilities include clarifying acceptable use, teaching verification, protecting

student information, and redesigning vulnerable assessments (Kumar, 2023; Maistry & Singh, 2025). These decisions are difficult when institutional rules lag behind changing tools or rely mainly on imperfect detection. Caution may therefore express stewardship rather than resistance. At the same time, governance focused only on prohibition can prevent faculty and students from developing the critical literacy needed to work responsibly with systems that are already entering academic and professional practice.

### *Instructional Continuity Versus Pedagogical Redesign*

Some uses make established tasks faster; others treat GenAI as an object of critique or a partner in inquiry. Meaningful integration may require faculty to reconsider what students should learn, how learning is demonstrated, and which judgments remain disciplinary responsibilities (Ellis et al., 2025; Kahn et al., 2025; Salhab, 2024). Faculty may ask students to compare generated and disciplinary explanations, document prompting and verification, or defend decisions that an output cannot justify for them. Such changes alter the distribution of epistemic work rather than simply adding a tool. The tension is therefore between incorporating GenAI into existing routines and redesigning learning around explicit purposes that remain educationally and disciplinarily defensible.

### *Technological Augmentation Versus the Human and Relational Meaning of Teaching*

Faculty also question how GenAI affects expertise, trust, dialogue, mentorship, and relationships with students (Farazouli et al., 2025; Mah & Groß, 2024). Teaching involves more than producing explanations or feedback; it also requires interpreting ambiguity, noticing development, and responding to students within ongoing relationships. Technology may support these responsibilities when saved time is redirected toward interaction or when generated alternatives deepen conversation. It may weaken them when automated outputs replace contextual judgment or make feedback appear personalized without reciprocal knowledge of the learner. This tension asks which human responsibilities should guide, rather than merely survive, technological use.

## Two Conceptual Dimensions

The tensions were organized along two analytical dimensions. The vertical dimension, generative possibility versus protective responsibility, distinguishes an emphasis on what GenAI can enable from an emphasis on what educators must safeguard. The horizontal dimension, operational practice versus pedagogical meaning and relationships, distinguishes attention to tasks, workflows, and rules from attention to learning, agency, professional purpose, and human relationships. These dimensions are not empirical scales or mutually exclusive positions; they clarify recurring emphases in the literature.

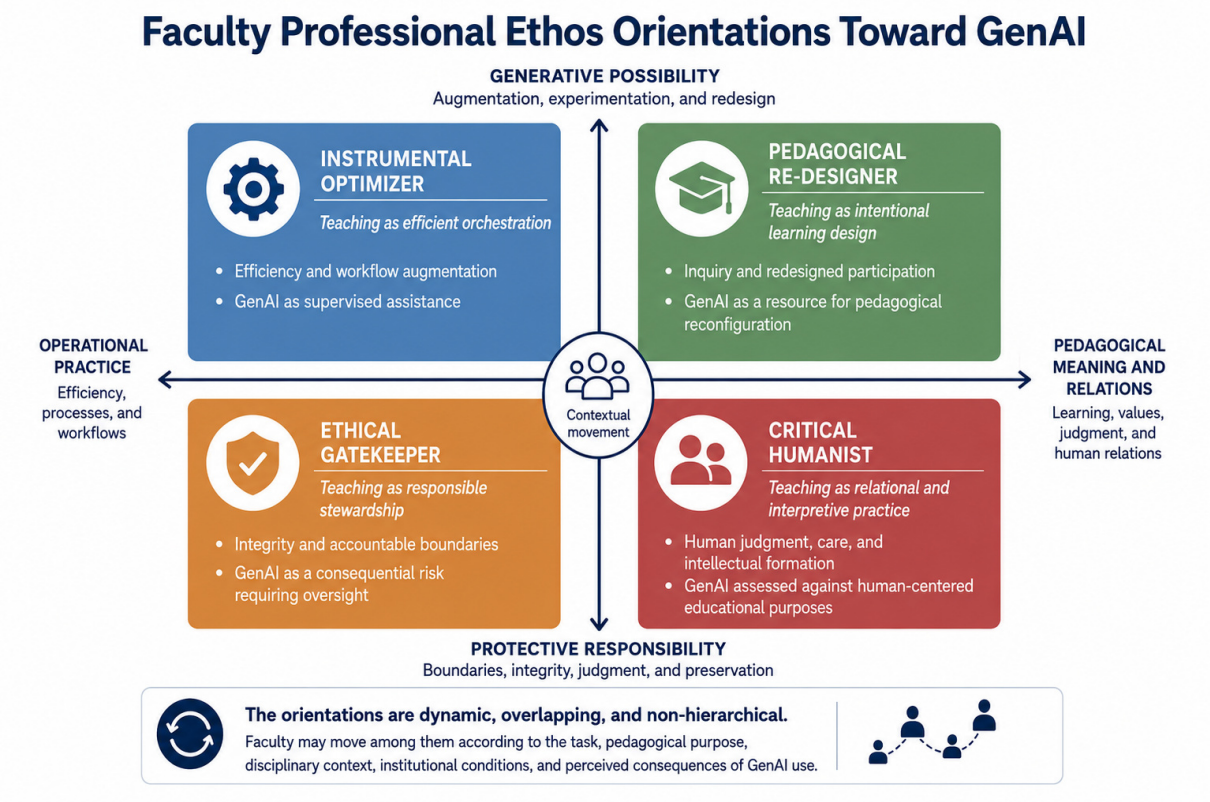
At the generative-potential end of the vertical dimension, faculty begin with experimentation, creativity, personalization, or expanded capacity. At the protective-responsibility end, they begin with integrity, privacy, validity, reliability, and retained human judgment. Protection does not equal rejection, just as possibility does not equal uncritical adoption. The distinction concerns which question receives priority: what can GenAI make possible, or what must responsible educators safeguard as it is used?

At the operational-practice end of the horizontal dimension, attention centers on tasks, procedures, workflows, policies, and immediate instructional performance. At the pedagogical-

meaning-and-relationships end, attention extends to the purposes of assessment, the distribution of agency, professional authority, mentorship, and the kinds of educational relationships a practice produces. Faculty can attend to both, but the distinction clarifies why an efficient or compliant practice may still prompt questions about learning and meaning.

Their intersection generates four orientations. Generative possibility combined with operational practice produces the Instrumental Optimizer. Protective responsibility combined with operational practice produces the Ethical Gatekeeper. Generative possibility combined with pedagogical meaning and relationships produces the Pedagogical Re-Designer. Protective responsibility combined with pedagogical meaning and relationships produces the Critical Humanist.

**Figure 1**
*Conceptual Typology of Faculty Professional Ethos Toward Generative AI in Higher Education*




The quadrants represent contextual emphases, not stable faculty types or a developmental sequence. One educator may optimize preparation, guard assessment boundaries, redesign an inquiry activity, and preserve human involvement in consequential feedback. The typology makes this movement analyzable without treating it as inconsistency.

## Four Orientations of Faculty Professional Ethos

### *The Instrumental Optimizer*

The Instrumental Optimizer treats GenAI primarily as a resource for extending faculty capacity. Its central commitment is effective practice: faster preparation, scalable feedback, resource generation, and reduced repetitive work (Agostini & Picasso, 2024; Hennekeuser et

al., 2024). Faculty authority becomes orchestrational: educators select tools, construct prompts, evaluate outputs, and decide how generated materials enter existing workflows. Ethical or pedagogical concerns may be acknowledged, but practical functionality remains the starting point. The orientation's limitation appears when efficiency is treated as sufficient evidence of educational value. Faster production may conceal new verification labor, normalize generic materials, or shift decisions whose consequences remain the faculty member's responsibility. Its central question is when assistance becomes delegation of professional judgment.

### *The Ethical Gatekeeper*

The Ethical Gatekeeper begins with stewardship. GenAI use is conditioned by integrity, fairness, transparency, privacy, reliability, and assessment validity (Ghimire et al., 2024; Kumar, 2023). Faculty clarify boundaries, protect information, verify outputs, and determine when disclosure or assessment redesign is required. Technology is permitted conditionally when oversight is meaningful and its use remains consistent with academic obligations. This orientation does not necessarily reject GenAI; its caution can protect trust at a time when authorship and evidentiary standards are unsettled. It can, however, become restrictive if responsibility is reduced to prohibition, surveillance, or detection. Its challenge is to maintain academic legitimacy while enabling informed participation and critical AI literacy.

### *The Pedagogical Re-Designer*

The Pedagogical Re-Designer treats GenAI as an occasion to reconsider learning design. GenAI becomes valuable when it supports inquiry, comparison, critique, iteration, collaboration, or AI literacy (Ellis et al., 2025; Lau & Guo, 2023; Salhab, 2024). Faculty act as architects of mediated learning: they specify purposes, design tasks, establish boundaries, and determine which judgments students must still make and defend. Authority is reconfigured through orchestration rather than ceded to a system. The orientation recognizes that students' access to GenAI changes what counts as a meaningful task or demonstration of learning. Its risk is mistaking technological novelty for pedagogical transformation. Redesign remains defensible only when it clarifies rather than displaces the knowledge, reasoning, and disciplinary participation students should develop.

### *The Critical Humanist*

The Critical Humanist foregrounds judgment, dialogue, mentorship, care, and intellectual presence. It asks what may be diminished when educational practices are automated and how GenAI affects agency, trust, and relationships (Farazouli et al., 2025; Mah & Groß, 2024; Pisica et al., 2023). Faculty remain responsible for interpreting ambiguity, understanding context, and responding to students as particular people rather than as inputs to a generalized system. This orientation does not require technological refusal. GenAI may serve human purposes when it creates time for dialogue or provides material for critical interpretation. Its limitation is the risk of idealizing existing human practices as automatically equitable, responsive, or sustainable. Its strongest contribution is to make relational purpose a criterion for technological judgment.

No orientation resolves every tension. Responsible engagement may combine the Optimizer's attention to capacity, the Gatekeeper's safeguards, the Re-Designer's pedagogical intentionality, and the Humanist's relational commitments. The framework therefore shifts analysis from whether faculty adopt GenAI to how they negotiate responsibility in particular practices.

## Discussion

Technology-adoption models explain important influences on faculty engagement, including usefulness, ease of use, social influence, and facilitating conditions (Davis, 1989; Venkatesh et al., 2003). The typology extends these accounts by explaining why acceptance and caution often coexist. Similar practices may express different commitments: restricting GenAI may protect assessment validity, privacy, foundational learning, or relational feedback, while encouraging it may pursue efficiency, disciplinary preparation, experimentation, or more time for human interaction.

### Professional Ethos as Dynamic and Contextual

The orientations are dynamic rather than hierarchical. This interpretation aligns with studies documenting distinct faculty profiles, qualitatively different teaching experiences, and context-responsive agency (Ellis et al., 2025; Kahn et al., 2025; Mah & Groß, 2024). Movement among orientations can represent differentiated judgment: faculty assess the responsibility implicated by a task and respond accordingly. GenAI does not produce one transformation of the faculty role; it makes assumptions about expertise, assessment, authorship, authority, and human presence available for reconsideration.

This dynamic account also explains why the orientations should not be ranked. The Pedagogical Re-Designer is not inherently more advanced than the Ethical Gatekeeper, and the Critical Humanist is not automatically more responsible than the Instrumental Optimizer. Each foregrounds a necessary commitment while leaving other questions less visible. Professional judgment lies partly in recognizing when another orientation must interrupt or supplement the first response to a task.

### Implications for Faculty Development and Policy

Faculty development focused only on technical proficiency may neglect when and why tools should be used, while risk-only training may suppress pedagogically valuable experimentation. Research instead supports combining technical competence with AI literacy, pedagogical judgment, ethical reflection, and collaborative learning (Cha et al., 2024; Cordero et al., 2025; Sjöberg & Bergdahl, 2025). The four orientations can structure reflection on: (a) tasks GenAI can support without transferring consequential judgment; (b) ethical and evidentiary boundaries; (c) needed redesign of learning and assessment; and (d) forms of human interpretation and relationship that should remain central.

Institutional policies should establish shared principles for transparency, privacy, fairness, accountability, and assessment validity while allowing disciplinary interpretation. Faculty require clear guidance but also professional discretion because authorship, evidence, acceptable assistance, and foundational knowledge differ across fields (Maistry & Singh, 2025; McGrath et al., 2023; Wainaina & Sun, 2025). Governance should support informed judgment through appropriate tools, sustained learning, and collective deliberation rather than shifting responsibility entirely to individual faculty or automated detection systems.

The framework can also support institutional dialogue without assigning faculty to types. A workshop might begin with a concrete teaching practice and ask how each orientation changes the questions posed about it. This makes disagreements more precise: participants can identify whether they are prioritizing capacity, stewardship, learning design, or relational purpose and

then examine what their preferred emphasis overlooks. Used in this way, the typology is a deliberative tool rather than an assessment of faculty readiness.

## Limitations and Future Research

This conceptual review was designed to assemble a theoretically relevant, not exhaustive or statistically representative, body of scholarship. Searches were limited to four discovery systems, citation tracking, and publications available in English through December 31, 2025. The 52 publications also differ in design and evidentiary purpose; empirical findings and conceptual claims were interpreted but not statistically aggregated. Placing 41 publications under one primary tension simplifies contributions that often span several tensions.

The orientations have not been validated as faculty profiles or measures, and the analysis does not establish their frequency or predictors. Future interview, classroom, policy, and longitudinal studies should examine how orientations combine and change across disciplines, institutions, and national contexts. The framework should initially function as a sensitizing typology open to revision, including additional dimensions or orientations, as GenAI and faculty practice evolve (Jaakkola, 2020).

Because GenAI capabilities and institutional rules are changing quickly, the December 2025 cutoff gives the synthesis temporal coherence but also limits its reach. Later research may foreground forms of labor, inequality, environmental cost, or governance that are less developed in the present corpus. Empirical studies should therefore test the framework against disconfirming cases rather than treating its four orientations as exhaustive.

## Conclusion

Faculty engagement with GenAI is not only a question of adoption. It involves judgments about capacity, integrity, learning, authority, and the human purposes of teaching. This synthesis organized four recurring tensions into four orientations: the Instrumental Optimizer, Ethical Gatekeeper, Pedagogical Re-Designer, and Critical Humanist.

The typology's contribution is not to classify educators but to articulate the commitments they negotiate. Responsible practice may require efficiency with accountability, safeguards with experimentation, redesign directed by learning, and technology guided by human relationships. For institutions, the framework supports professional learning and policies beyond technical training or generalized prohibition. For research, it provides a vocabulary for examining what responsible teaching becomes when generative systems participate in educational work.

The question facing higher education is thus not simply whether faculty will adopt GenAI. It is how educators and institutions will decide what may be delegated, what must be redesigned, and what forms of judgment and relationship should remain constitutive of teaching.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Grammarly, an AI-assisted writing software, was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The author also declares that, apart from Grammarly, ChatGPT (OpenAI) have been used to assist with

formatting the literature-synthesis table and creating the visual design of the conceptual typology figure based on categories, dimensions, and interpretations developed and specified by the author. The author further declares that no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Agostini, D., & Picasso, F. (2024). Large language models for sustainable assessment and feedback in higher education. *Intelligenza Artificiale, 18*, 121–138. https://doi.org/10.3233/IA-240033

Al-Dosari, I. K. (2023). The role of artificial intelligence in the self-professional development of faculty members: Prince Sattam bin Abdulaziz University as a model. *ArtsEduca, 37*, 260–277. https://doi.org/10.58262/ArtsEduca.3720

Al-Mughairi, H., & Bhaskar, P. (2023). Exploring the factors affecting the adoption of AI techniques in higher education: Insights from teachers' perspectives on ChatGPT. *Journal of Research in Innovative Teaching & Learning, 18*(2), 232–247. https://doi.org/10.1108/JRIT-09-2023-0129

Alammari, A. (2024). Evaluating generative AI integration in Saudi Arabian education: A mixed-methods study. *PeerJ Computer Science, 10*, e1879. https://doi.org/10.7717/peerj-cs.1879

Ali, H. O. (2024). University teachers' vantage points on ChatGPT integration in education: Upsides and downsides. *Rotura – Revista de Comunicação, Cultura e Artes, 4*(1).

Alrayes, A., Henari, T. F., & Dalal, A. A. (2024). ChatGPT in education—Understanding the Bahraini academics' perspective. *Electronic Journal of e-Learning, 22*(2), 112–134. https://doi.org/10.34190/ejel.22.2.3250

Barrett, A., & Pack, A. (2023). Not quite eye to A.I.: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education, 2*0, 59. https://doi.org/10.1186/s41239-023-00427-0

Cha, Y., Dai, Y., Lin, Z., Liu, A., & Lim, C. P. (2024). Empowering university educators to support generative AI-enabled learning: Proposing a competency framework. *Procedia CIRP, 128*, 256–261. https://doi.org/10.1016/j.procir.2024.06.021

Cordero, J., Torres-Zambrano, J., & Cordero-Castillo, A. (2025). Integration of generative artificial intelligence in higher education: Best practices. *Education Sciences, 15*(1), 32. https://doi.org/10.3390/educsci15010032

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

de Obesso, M. M., Núñez-Canal, M., Pérez-Rivero, C. A., & Álvarez-de-Mon, I. (2025). University professors' level of performance in the use of artificial intelligence. *Journal of Competitiveness, 17*(4), 349–377. https://doi.org/10.7441/joc.2025.04.13

Deroncele-Acosta, A., Sayán-Rivera, R. M. E., Mendoza-López, A. D., & Norabuena-Figueroa, E. D. (2025). Generative artificial intelligence and transversal competencies in higher education: A systematic review. *Applied System Innovation, 8*(3), 83. https://doi.org/10.3390/asi8030083

Ellis, R., Han, F., & Cook, H. (2025). Qualitatively different teacher experiences of teaching with generative artificial intelligence. *International Journal of Educational Technology in Higher Education, 22,* 33. https://doi.org/10.1186/s41239-025-00532-2

Farazouli, A., Cerratto Pargman, T., Bolander Laksov, K., & McGrath, C. (2025). Navigating uncertainty: University teachers' experiences and perceptions of generative artificial intelligence in teaching and learning. *Studies in Higher Education*. Advance online publication. https://doi.org/10.1080/03075079.2025.2550766

Gaber, S. A., Shahat, H. A., Alkhateeb, I. A., Al Hasan, S. A., Alqatam, M. A., Almughyirah, S. M., & Kamel, M. K. (2023). Faculty members' awareness of artificial intelligence and its relationship to technology acceptance and digital competencies at King Faisal University. *International Journal of Learning, Teaching and Educational Research, 22*(7), 473–496. https://doi.org/10.26803/ijlter.22.7.25

Ghimire, A., Prather, J., & Edwards, J. (2024). Generative AI in education: A study of educators' awareness, sentiments, and influencing factors. *Journal of Educational Technology & Society, 27*(1), 154–173. https://doi.org/10.48550/arxiv.2403.15586

Gong, D., & Xiao, P. (2024). Research on the enhancement of informatization literacy of design practice teaching faculty in applied colleges and universities based on OBE model in the context of artificial intelligence. *Applied Mathematics and Nonlinear Sciences, 9*(1), 1–17. https://doi.org/10.2478/amns-2024-0971

Hasan, H. E., Jaber, D., Al Tabbah, S., Lawand, N., Habib, H. A., & Farahat, N. M. (2024). Knowledge, attitude, and practice among pharmacy students and faculty members towards artificial intelligence in pharmacy practice: A multinational cross-sectional study. *PLOS ONE, 19*(3), e0296884. https://doi.org/10.1371/journal.pone.0296884

Hennekeuser, D., Vaziri, D. D., Golchinfar, D., Schreiber, D., & Stevens, G. (2024). Enlarged education—Exploring the use of generative AI to support lecturing in higher education. *International Journal of Artificial Intelligence in Education, 35*(3), 1096–1128. https://doi.org/10.1007/s40593-024-00424-y

Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review, 10*, 18–26. https://doi.org/10.1007/s13162-020-00161-0

Jose, J., & Jose, B. J. (2024). Educators' academic insights on artificial intelligence: Challenges and opportunities. *Electronic Journal of e-Learning, 22*(2), 59–77. https://doi.org/10.34190/ejel.21.5.3272

Kahn, P., Carrigan, M., Smith, P., Murtagh, L., Liu, R., & Song, F. (2025). Teacher agency and generative artificial intelligence: Teaching in higher education as a responsive, cultural activity. *Learning, Media and Technology*. Advance online publication. https://doi.org/10.1080/17439884.2025.2575993

Kennedy, M. M. (2007). Defining a literature. *Educational Researcher, 36*(3), 139–147. https://doi.org/10.3102/0013189X07299197

Kiryakova, G., & Angelova, N. (2023). ChatGPT—A challenging tool for the university professors in their teaching practice. *Education Sciences, 13*(10), 1056. https://doi.org/10.3390/educsci13101056

Kumar, R. (2023). Faculty members' use of artificial intelligence to grade student papers: A case of implications. *International Journal for Educational Integrity, 19*, 9. https://doi.org/10.1007/s40979-023-00130-7

Lau, S., & Guo, P. J. (2023). From "ban it till we understand it" to "resistance is futile": How university programming instructors plan to adapt as more students use AI code generation and explanation tools such as ChatGPT and GitHub Copilot. *Proceedings of the 2023 ACM Conference on International Computing Education Research*, 106–121. https://doi.org/10.1145/3568813.3600138

Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F., Lekkas, D., & Palmer, E. (2024). The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers and Education: Artificial Intelligence, 6*, 100221. https://doi.org/10.1016/j.caeai.2024.100221

Mah, D.-K., & Groß, N. (2024). Artificial intelligence in higher education: Exploring faculty use, self-efficacy, distinct profiles, and professional development needs. *International Journal of Educational Technology in Higher Education, 21*, 58. https://doi.org/10.1186/s41239-024-00490-1

Maistry, S. M., & Singh, U. G. (2025). Faculty perspectives on the role of ChatGPT-4.0 in higher education assessments. *Journal of Education, 98*, 86–102. https://doi.org/10.17159/2520-9868/i98a05

Marshall, C., Rossman, G. B., & Blanco, G. L. (2022). *Designing qualitative research* (7th ed.). SAGE.

McGrath, C., Cerratto Pargman, T., Juth, N., & Palmgren, P. J. (2023). University teachers' perceptions of responsibility and artificial intelligence in higher education: An experimental philosophical study. *Computers and Education: Artificial Intelligence, 4*, 100139. https://doi.org/10.1016/j.caeai.2023.100139

Miranda, F. J., & Chamorro-Mera, A. (2026). Exploring the adoption of generative artificial intelligence tools among university teachers. *Higher Education Research & Development, 45*(3), 680–696. https://doi.org/10.1080/07294360.2025.2559648

Nikoçeviq-Kurti, E., & Bërdynaj-Syla, L. (2024). ChatGPT integration in higher education: Impacts on teaching and professional development of university professors. *Education Process: International Journal, 13*(3), 22–39. https://doi.org/10.22521/edupij.2024.133.2

Ofosu-Ampong, K. (2024). Beyond the hype: Exploring faculty perceptions and acceptability of AI in teaching practices. *Discover Education, 3*, 38. https://doi.org/10.1007/s44217-024-00128-4

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence, 2*, 100020. https://doi.org/10.1016/j.caeai.2021.100020

Pang, T. Y., Kootsookos, A., & Cheng, C.-T. (2024). Artificial intelligence use in feedback: A qualitative analysis. *Journal of University Teaching & Learning Practice, 21*(6). https://doi.org/10.53761/40wmcj98

Pisica, A. I., Edu, T., Zaharia, R. M., & Zaharia, R. (2023). Implementing artificial intelligence in higher education: Pros and cons from the perspectives of academics. *Societies, 13*(5), 118. https://doi.org/10.3390/soc13050118

Rajabi, P., Taghipour, P., Cukierman, D., & Doleck, T. (2023). Exploring ChatGPT's impact on post-secondary education: A qualitative study. *Proceedings of the 25th Western Canadian Conference on Computing Education*, 1–6. https://doi.org/10.1145/3593342.3593360

Salhab, R. (2024). AI literacy across curriculum design: Investigating college instructors' perspectives. *Online Learning, 28*(2), 22–47. https://doi.org/10.24059/olj.v28i2.4426

Simelane, P. M., & Kittur, J. (2025). Use of generative artificial intelligence in teaching and learning: Engineering instructors' perspectives. *Computer Applications in Engineering Education, 33*(1), e22813. https://doi.org/10.1002/cae.22813

Sjöberg, J., & Bergdahl, N. (2025). University teachers' reflections on the use of generative AI in higher education: An exploratory study. In B. K. Smith & M. Borge (Eds.), *HCII 2025: Lecture Notes in Computer Science* (Vol. 15807, pp. 173–186). Springer. https://doi.org/10.1007/978-3-031-93567-1_12

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Soodan, V., Rana, A., Jain, A., & Sharma, D. (2024). AI chatbot adoption in academia: Task fit, usefulness, and collegial ties. *Journal of Information Technology Education: Innovations in Practice, 23,* 1–25. https://doi.org/10.28945/5260

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review, 4*(3), 356–367. https://doi.org/10.1177/1534484305278283

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

Verano-Tacoronte, D., Bolívar-Cruz, A., & Sosa-Cabrera, S. (2025). Are university teachers ready for generative artificial intelligence? Unpacking faculty anxiety in the ChatGPT era. *Education and Information Technologies, 30*, 20495–20522. https://doi.org/10.1007/s10639-025-13585-7

Wainaina, P. K., & Sun, Y. (2025). Educators' perceptions and willingness to integrate generative artificial intelligence in teaching and research: Evidence from Kenyan higher education. *Discover Education, 4*, 347. https://doi.org/10.1007/s44217-025-00820-z

Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education, 21*, 21. https://doi.org/10.1186/s41239-024-00453-6

**Contact email:** agamakru@u.rochester.edu

# Implementation of Integrated Curriculum in the Opinions of Primary Education Teachers

Monika Brozmanová, Matej Bel University Banská Bystrica, Slovakia
Ivana Rochovská, Matej Bel University Banská Bystrica, Slovakia

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

The integration of educational content is one of the conceptual pillars of the current curricular reform in primary education in Slovakia. It represents a shift from a fragmented subject-based approach to meaningful, interconnected learning that contributes to the comprehensive development of students' competences. The implementation of an integrated curriculum in primary education is not just an organisational change, but also requires a change in teachers' professional thinking. Its practical implementation depends on teachers' professional understanding of changes in the curriculum and their project competences. The aim of this paper is to analyse how primary school teachers conceptualise the integration of educational content, what their experiences are, how they approach their own design of an integrated curriculum, and what barriers they identify in this process. The research was conducted using a qualitative design and focus group interviews. The data were processed by thematic content analysis focusing on semantic frameworks, professional strategies, and systemic limits to the implementation of reform principles. The findings suggest that the participating teachers perceive integration as a pedagogically meaningful approach that supports contextualised student learning, but its implementation is conditioned by the degree of curricular autonomy, time available for planning, and the level of methodological support. Their statements reveal a tension between the declared flexibility of curricular documents and the practical reality of schools. The results of the study contribute to a deeper understanding of the implementation of curriculum reform at the level of pedagogical practice and point to the need for systematic support for teachers in the creation and implementation of an integrated curriculum, including adequate initial teacher education and continuing professional development. The study emphasises the importance of supporting teachers in the areas of interdisciplinary project design, ensuring content coherence, and reflective professional development.

*Keywords:* integration of educational content, curriculum reform, key competences, primary school teacher, focus group interview

iafor
The International Academic Forum
www.iafor.org

## Introduction

Current curricular reforms across many educational systems are shifting from the fragmented acquisition of isolated knowledge towards fostering students' holistic development, competence development, and meaningful connections across educational domains. The curricular reform introduced in Slovak primary education in 2023 reflects this international trend by promoting an integrated approach to curriculum design and implementation. A key principle of the reform is the integration of educational content, based on the premise that learners construct knowledge through interconnected experiences rather than within the boundaries of isolated school subjects.

The *State Educational Programme for Primary Education* (Ministry of Education, Research, Development and Youth of the Slovak Republic & National Institute of Education and Youth, 2023) represents a substantial transformation in the goals and organisation of primary education. The reform emphasises contextualised learning, interdisciplinary connections, and the development of competences required for life in the twenty-first century. Its objectives include supporting students' holistic development, fostering critical thinking and problem-solving skills, strengthening communication and collaboration, promoting learner autonomy and responsibility, encouraging creativity and initiative, creating an inclusive and stimulating learning environment, connecting learning with authentic real-life situations, and developing civic, social, and digital competences.

These objectives have important implications for curriculum design. Within the reform framework, content integration is understood not merely as combining topics from different school subjects but as the purposeful creation of meaningful relationships among educational objectives, learning content, and students' authentic experiences. Such an approach is expected to enhance conceptual understanding, facilitate knowledge transfer, and promote the development of higher-order thinking skills.

The successful implementation of an integrated curriculum, however, depends largely on teachers' professional understanding of the concept and their ability to translate curricular intentions into classroom practice. Designing integrated learning experiences requires teachers to identify meaningful relationships among educational standards, formulate interdisciplinary learning objectives, and create coherent learning activities that support competence development.

Against this background, the present study aims to explore how primary school teachers conceptualise content integration, examine their experiences with its implementation, investigate the ways in which they design an integrated curriculum, and identify the barriers they encounter during this process.

## Theoretical Foundations of Curriculum Integration

Curriculum integration is not a new topic in educational research. Constructivist theories of learning have long emphasised the importance of connecting knowledge and actively constructing meaning through experience (Piaget, 1970; Vygotsky, 1978). In the Czech and Slovak pedagogical tradition, Chlup (1957) was among the first to warn against the didactic limitations of fragmented knowledge acquisition and advocated a dialectical approach to teaching that encourages thinking in relationships and broader contexts. Similarly, Maňák (2024) argues that neglecting interdisciplinary connections leads to isolated knowledge

structures, whereas instruction should promote the synthesis of knowledge and skills across subject boundaries.

Integrated thematic instruction represents a pedagogical approach that connects knowledge from different educational domains through common themes or authentic problems, thereby creating more meaningful learning experiences. Kerry (2015) suggests that cross-curricular teaching enables teachers to design more creative learning situations while fostering deeper understanding. Likewise, Björklund and Ahlskog-Björkman (2017) demonstrated that integrating mathematics and the arts in early childhood education promotes holistic learning and supports children's ability to establish conceptual connections from the beginning of formal education.

Evidence from cognitive psychology further supports integrated curricula. Bransford et al. (2000) argue that meaningful learning occurs when learners recognise relationships among concepts and transfer knowledge to new contexts. Similarly, Perkins (2009) emphasises the importance of transfer learning, whereby knowledge acquired in one domain can be effectively applied in different situations. Curriculum integration therefore facilitates the development of interconnected cognitive structures and enhances conceptual understanding.

Contemporary neuroscience provides additional theoretical support for integrated learning. Brain research indicates that learning is mediated through interconnected neural networks rather than isolated subject-specific systems (Sousa, 2022). Jensen (2008) argues that learning is more effective when new information is embedded in meaningful contexts and connected with prior knowledge, emotions, experiences, and social interaction. From this perspective, integrated thematic instruction aligns with natural learning processes by promoting multiple representations of knowledge and meaningful cognitive connections.

The rationale for curriculum integration is also grounded in developmental psychology. Young learners perceive the world as an interconnected whole rather than as separate school subjects. Consequently, integrated learning enables students to better understand relationships among phenomena, transfer knowledge across contexts, and develop higher-order thinking skills.

The effectiveness of integrated curricula has been supported by several intervention studies conducted in educational settings (Portaankorva-Koivisto & Havinga, 2019; Ratcliffe et al., 2005; Siew & Ambo, 2018; Šindelková et al., 2016). However, successful implementation depends not only on curriculum design but also on teachers' understanding and enactment of reform principles. Fullan (2007) emphasises that teachers are the key agents of educational change and that sustainable reform occurs only when they understand and internalise its underlying purposes.

Despite the documented benefits of curriculum integration, empirical studies also identify significant implementation challenges. Effective integrated teaching requires teachers to possess a deep understanding of disciplinary content as well as the ability to identify and intentionally construct meaningful interdisciplinary connections (Cherni, 2020). Furthermore, integrated curriculum planning places considerable demands on teachers' pedagogical design, instructional planning, and selection of appropriate teaching strategies (Yunisrul et al., 2021). Successful implementation also depends on collaboration and coordinated planning among teachers, which remains one of the most challenging organisational aspects of curriculum integration in school practice (Retnawati et al., 2018).

## Methods

A qualitative research design was employed to explore primary school teachers' perceptions of curriculum integration. Data were collected through semi-structured focus group interviews, which served as a pilot study within a broader research project investigating teachers' approaches to curriculum integration in the context of the Slovak curricular reform.

The interview protocol combined predetermined questions with flexible follow-up prompts, enabling participants to elaborate on their experiences and perspectives. The discussions focused on teachers' understanding of curriculum integration, their previous experiences with integrated teaching, approaches to designing integrated curricula, perceived barriers to implementation, and their evaluation of concept maps as a planning tool. All interviews were audio-recorded, transcribed verbatim, and analysed using inductive open coding.

The pilot sample consisted of four female primary school teachers. Three participants worked in schools that had not joined the curricular reform pilot programme and had between three and seven years of teaching experience. One participant taught in a school involved in the second year of the reform pilot and had five years of teaching experience.

At the beginning of each interview, participants were introduced to the Integrated Thematic Instruction model (Kovalik & Olsen, 1996) and to a set of concept maps developed by the research team. The concept maps were created through qualitative content analysis of the educational standards for individual curriculum areas. Units of analysis included words, phrases, and sentences representing concepts, themes, relationships, facts, or activities. The coding process was conducted independently by two experts—a specialist in primary education and a subject-specific didactics expert—to enhance analytical reliability. Differences in coding were subsequently discussed until consensus was reached, and the agreed categories formed the basis for constructing the concept maps used during the interviews.

## Results of the Focus Group Interviews

The results of the content analysis of the focus group interviews are presented in Table 1. Segments of the interview transcripts were assigned codes, which are listed in the right-hand column, and were grouped into thematic categories (left-hand column).

**Table 1**
*Results of the Coding From the Content Analysis of the Interviews*

| Category | Code |
|---|---|
| Understanding the concept of content integration | Integrated projects |
| | Theme day |
| | Interdisciplinary links |
| | Inclusion |
| | Incorrect or narrow understanding of the concept |
| | Lack of knowledge of the theoretical foundations of content integration |
| | Perception of a lack of knowledge of content integration amongst colleagues |
| | Positive self-assessment of understanding of content integration |

| | |
|---|---|
| Designing an integrated curriculum | Intuitive planning<br>Working with the curriculum is largely implicit<br>Flexibility in terms of time and organisation<br>Combining resources<br>The curriculum as a starting point for lesson planning<br>Teaching based on pupils' interests<br>The need for professional development and participation in training courses |
| Views on concept maps | Clearer guidance on content standards<br>Positive feedback – helpful for teachers |
| The benefits of introducing integrated teaching | Freedom to choose topics to meet the standards<br>Joint decision-making on the choice of topics<br>Greater teacher autonomy |
| Barriers to integrated teaching | Topics set by the school management<br>Discrepancies between the content of different subjects within the planned curriculum (in the case of a project or theme day)<br>The need for further training in the integration of educational content |

*Source:* Compiled by the author

## Teachers' Perceptions of Content Integration

The results suggest that the teachers participating in this pilot study tend to view content integration in a rather narrow and pragmatic way. Most often, they equate it with organising themed and project days (e.g. Apple Day, Earth Day, Tree Day) or with linking the same topics across different subjects. Whilst the teachers do create thematic coherence for the day, they lack an explicit link to common educational objectives and a conceptual connection between content standards.

## Theoretical Foundations of Content Integration

The results point to conceptual ambiguity, as some teachers also associate the term "integration" with the inclusion of pupils with special educational needs.

The interviews revealed a lack of theoretical grounding regarding the issue of content integration. Terms such as "integrated thematic teaching," "brain-compatible learning" or the names of authors such as S. Kovalik and K. Olsen are, according to the teachers, completely unfamiliar to them. The teachers reflected on the lack of systematic preparation for the implementation of the reformed curriculum and pointed out that the training they had undergone had focused primarily on a mechanical approach to working with educational standards. At the same time, they expressed an interest in further professional development that would foster a deeper understanding of the principles of content-integrated teaching.

## Designing an Integrated Curriculum

Teachers described the design of an integrated curriculum as a largely intuitive process based on pedagogical experience, textbooks and pupils' interests.

Although they stated that they focused on integrating content within the educational standards, the use of these standards in lesson planning remains largely implicit. The interviews also revealed a perception that teaching is focused on the pupil and their needs, whilst emphasising the diversity of classes and the need to adapt teaching to the specific class group. They view content integration as a means of promoting motivation, active learning and respect for pupils' individual needs.

### Perceptions of Concept Maps

Concept maps were assessed by the teachers as an effective methodological aid in designing an integrated curriculum. They appreciated their contribution to navigating educational standards, identifying content connections and exploring new ways of linking educational content (e.g. linking tourism within the social sciences curriculum to physical education). The results suggest their potential to support more systematic and conscious planning of integrated teaching.

### Barriers and Prospects for Content Integration

The freedom to choose topics and support from the school management were cited as the most significant positive factors in the context of the discussion on the integration of educational content. Where the teachers have the opportunity to have a say in choosing topics, they perceive content integration as more meaningful and authentic. Their visions point towards greater autonomy, systematic methodological support and continuous professional development in the field of content-integrated teaching.

The teachers identified the main obstacles to the implementation of integrated teaching as limited autonomy in the choice of topics – which is linked to regulations imposed by the school management – inconsistencies between the content of individual subjects (in the case of project-based or thematic days) and insufficient methodological support. These obstacles lead to a sense of uncertainty and confusion when seeking suitable links, which in turn highlights the need for clear methodological tools and support within the curriculum.

## Discussion

The results of the qualitative study suggest that, among the participating primary school teachers, content integration has not yet been clearly conceptualised despite the ongoing implementation of the reformed curriculum. Most participants interpret it in terms of thematic or project days, or as linking the same topics across different subjects. However, this understanding represents the realisation of cross-curricular links or the thematic contextualisation of the curriculum rather than content integration based on the conscious linking of educational objectives, content standards and expected learning outcomes. The results therefore suggest a discrepancy between the theoretical concept of content integration presented in the reformed curriculum and its interpretation in pedagogical practice.

This finding is consistent with research indicating that the implementation of an integrated curriculum requires a significantly higher level of curricular thinking on the part of the teacher than traditional lesson planning (Cherni, 2020). The integration of educational content involves not only linking subject areas thematically, but above all designing shared learning objectives and deliberately establishing connections between knowledge from different subject areas. The

results of the interviews suggest that this shift in the participating teachers' professional thinking is not yet sufficiently established.

Another significant finding is the acknowledged lack of theoretical knowledge regarding content integration. Teachers openly admit that their professional training for the implementation of the reformed curriculum focused predominantly on working with educational standards, whilst the principles of content integration themselves took a back seat. These findings support the conclusions of Yunisrula et al. (2021), according to whom designing an integrated curriculum is one of the most professionally demanding tasks for a teacher. Similarly, Retnawati et al. (2018) point out that the successful implementation of integrated teaching requires not only the teacher's individual competences, but also systematic methodological support and coordination of lesson planning.

The results also showed that the participating teachers intuitively apply several principles of content integration, particularly focusing on pupils' interests, creating authentic learning situations and linking the curriculum to everyday experiences. These principles are consistent with constructivist approaches to learning; however, they are often put into practice without conscious curriculum design. It can therefore be assumed that the problem with implementation does not lie in the absence of integration activities, but rather in their insufficient theoretical grounding and systematic planning.

The teachers themselves highlight the difficulty of transitioning from a transmissive model of teaching to student-centred learning, which is one of the fundamental requirements of the current curriculum. This aspect confirms that the implementation of the reform is not primarily a matter of adopting new methodological approaches, but of changing professional beliefs and pedagogical thinking.

The evaluation of concept maps can be considered particularly insightful. The teachers unequivocally identified them as an effective methodological aid in designing an integrated curriculum. In particular, they praised their ability to visualise relationships between learning content, to facilitate navigation of the educational standards, and to encourage the exploration of new ways of linking subject matter. These findings suggest that one effective tool for supporting the implementation of the reformed curriculum may be the creation of methodological materials that present content relationships to teachers in a clear, visual format.

Another significant finding is the discrepancy between the participating teachers' positive attitudes towards content integration and the conditions under which it is implemented. Although they regard integration as beneficial in terms of pupil motivation and the authenticity of learning, they highlight their limited autonomy in choosing topics, the time-consuming nature of planning, and a lack of methodological support. These findings suggest that the successful implementation of an integrated curriculum does not depend solely on teachers' individual competences, but is also conditioned by the school's systemic conditions, professional development and support from the school management.

## Summary and Recommendations

The implementation of curricular changes begins in Year 1 of primary school – that is, in primary education – and the guiding principles of the education reform are gradually being incorporated into and shaping higher levels of education as well. In the reformed curriculum, the role of the teacher is changing most significantly, as teachers are becoming curriculum

designers: under the two-tier curriculum, each school develops its own school curriculum based on the national curriculum.

The findings suggest that content integration is still in a phase of gradual adoption among the participating teachers. They express a positive attitude towards the reform changes, are highly motivated to develop their professional competences, and recognise the need for further professional development. Future professional development should therefore be directed not only towards the acquisition of methodological approaches, but above all towards a deeper understanding of the theoretical principles of a content-integrated curriculum and their application in the systematic planning of lessons.

The reform moves away from the traditional emphasis on the isolated acquisition of knowledge in individual subjects and focuses on developing pupils' comprehensive competences by linking knowledge, skills, attitudes and experiences from different areas of education. The new curriculum creates scope for linking educational areas through thematic blocks, project-based teaching, problem-based learning and research activities. It can be said that the integration of educational content represents one of the key requirements of the reformed curriculum. This is not merely an organisational change in teaching, but a shift in the philosophy of education, which should, above all, be embraced by teachers – the primary and most important agents of change in education.

## Acknowledgements

Both this article and the authors' participation in the international conference are outputs of the VEGA grant project No. 1/0486/24, "'Research into teachers' potential and analysis of curriculum documents from the perspective of the integration of educational content in primary education."

## References

Björklund, C., & Ahlskog-Björkman, E. (2017). Approaches to teaching in thematic work: Early childhood teachers' integration of mathematics and art. *International Journal of Early Years Education, 25*(2), 98–111. https://doi.org/10.1080/09669760.2017.1287061

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school* (Expanded ed.). National Academy Press. https://doi.org/10.17226/9853

Cherni, T. (2020). Analysis of the benefits and drawbacks of a thematic approach to designing an innovative curriculum [Unpublished manuscript]. https://www.researchgate.net/publication/352296505_Analysis_of_the_Benefits_and_Drawbacks_of_a_Thematic_Approach_to_Designing_an_Innovative_Curriculum

Chlup, O. (1957). *K základním otázkám pedagogiky: Výbor z díla* [On fundamental issues of pedagogy: Selected works]. Státní pedagogické nakladatelství.

Fullan, M. (2007). *The new meaning of educational change* (4th ed.). Teachers College Press.

Jensen, E. (2008). *Brain-based learning: The new paradigm of teaching* (2nd ed.). Corwin Press.

Kerry, T. (Ed.). (2015). *Cross-curricular teaching in the primary school: Planning and facilitating imaginative lessons* (2nd ed.). Routledge.

Maňák, J. (2024). *Nárys didaktiky* [Outline of didactics]. Masaryk University.

Ministry of Education, Research, Development and Youth of the Slovak Republic, & National Institute for Education and Youth. (2023). *National curriculum for primary education*. https://www.minedu.sk/data/files/11808_statny-vzdelavaci-program-pre-zakladne-vzdelavanie-cely.pdf

Perkins, D. (2009). *Making learning whole: How seven principles of teaching can transform education.* Jossey-Bass.

Piaget, J. (1970). *Psychology and pedagogy.* Viking Press.

Portaankorva-Koivisto, P., & Havinga, M. (2019). Integrative phenomena in visual arts and mathematics. *Journal of Mathematics and the Arts, 13*(1–2), 4–24. https://doi.org/10.1080/17513472.2018.1504269

Ratcliffe, M., Harris, R., & McWhirter, J. (2005). Cross-curricular collaboration in teaching social aspects of genetics. In K. Boersma, M. Goedhart, O. de Jong, & H. Eijkelhof (Eds.), *Research and the quality of science education* (pp. 77–88). Springer. https://doi.org/10.1007/1-4020-3673-6_7

Retnawati, H., Djidu, H., Kartianom, Apino, E., & Anazifa, R. D. (2018). Teachers' knowledge about higher-order thinking skills and its learning strategy. *Problems of Education in the 21st Century, 76*(2), 215–230. https://doi.org/10.33225/pec/18.76.215

Siew, N. M., & Ambo, N. (2018). Development and evaluation of an integrated project-based and STEM teaching and learning module on enhancing scientific creativity among fifth graders. *Journal of Baltic Science Education, 17*(6), 1017–1033. https://doi.org/10.33225/jbse/18.17.1017

Sousa, D. A. (2022). *How the brain learns* (6th ed.). Corwin Press.

Šindelková, M., Plucková, I., & Málková, A. (2016). Project-based education or integrated thematic instruction. In *13th International Student Conference on Project-Based Education in Science Education* (pp. 148–153). Charles University in Prague, Faculty of Education.

Yunisrul, Y., Zuardi, Z., & Arwin, A. (2021). Improvement of integrated thematic learning process using problem-based learning model. In *Proceedings of the First International Conference on Economics, Business and Social Humanities (ICONEBS 2020).* EAI. https://doi.org/10.4108/eai.4-11-2020.2304593

**Contact emails:** monika.brozmanova@umb.sk
ivana.rochovska@umb.sk

# Designing for Disciplinary Diversity: A Comparative Analysis of Preferences in Informal Learning Environments

Siriwan Rujibhong, Bangkok University, Thailand

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

In the contemporary academic landscape, Informal Learning Environments (ILEs) serve as critical infrastructure for fostering productivity, cross-disciplinary inspiration, and social cohesion. This research investigates the inherent nexus between disciplinary nature and spatial requirements through a detailed comparative study of students within Creative and STEM (Science, Technology, Engineering, and Mathematics) clusters. The study addresses a central problem: how the distinct spatial behaviours and co-cognitive demands of different academic disciplines necessitate specialised interior design interventions within shared university territories. Employing an Exploratory Sequential Mixed-Methods design, the research initially utilised qualitative in-depth interviews to map the nuanced environmental preferences and habitus of both cohorts. These qualitative insights informed the subsequent development of a quantitative survey administered to a sample of 305 students. Data analysis validates significant distinctions: whilst Creative students demonstrate a clear proclivity for high-modularity environments and collaborative-intensive zones that facilitate visual stimuli, STEM students prioritise acoustic insulation, technological integration, and focus-intensive configurations optimised for sustained cognitive work. The findings reveal that students from different disciplines have distinct requirements and exhibit unique spatial characteristics rooted in their respective academic backgrounds. Consequently, the research argues against homogenous universalist design, advocating instead for flexibility, adjustable elements, and the implementation of various specialised zones. The study culminates in strategic recommendations for hybrid configurations that favour a balance between the fluid, collaborative needs of the creative process and the precision-oriented requirements of STEM disciplines.

*Keywords:* informal learning environments, third place, learning space design, disciplinary culture, mixed methods, higher education

iafor
The International Academic Forum
www.iafor.org

## Introduction

Higher education institutions have, particularly since the early 2000s, substantially expanded their allocation of resources to spaces that sit outside the traditional lecture theatre or seminar room. Co-working lounges, informal study commons, maker spaces, and campus cafés are now routinely positioned as essential infrastructure for student learning, socialisation, and wellbeing, rather than as incidental leftover space between classrooms. These settings are referred to collectively as informal learning environments (ILEs). While the broader distinction between formal and informal learning dates to the mid-twentieth century, the dedicated ILE first appeared on university campuses in the early-to-mid 1990s in the form of the library "information commons," which reorganised study space around shared access to emerging digital resources (Beagle, 1999). The Information Arcade at the University of Iowa (1992) and the Information Commons at the University of Southern California (1994) are among the earliest and most widely cited examples (Heitsch & Holley, 2011). Over the following two decades, this model evolved from the technology-centred information commons into the more socially oriented "learning commons" and, ultimately, into a broader re-conception of the entire campus as a learning space (Jamieson, 2009). This trajectory reflects a growing recognition within the learning-space literature that a significant proportion of meaningful student learning, collaboration, and identity formation occurs outside scheduled, instructor-led activity, in spaces that students choose for themselves, on their own terms.

### Research Problem

Despite growing institutional investment in informal learning environments, design decisions are frequently driven by aesthetic trends or generic templates rather than empirical evidence of how students utilise space. This standardised planning conflicts with the foundational premise that different academic fields possess unique learning traits and behavioural needs. As established by Biglan (1973), disciplines occupy distinct knowledge domains. Consequently, students approach their work through separate cognitive styles grounded in Kolb's (1984) experiential learning theory (Biglan, 1973; Kolb, 1984).

These distinct profiles create disparate spatial requirements. For example, Fine Arts students require informal spaces designed to store and display in-progress physical artefacts safely. Conversely, Engineering students require environments optimised for testing and assembling electronic prototypes. Because a generic design ignores these distinct academic tasks, it risks satisfying neither group. This represents both a practical institutional inefficiency and a missed opportunity to leverage physical space as a strategic tool for supporting disciplinary identity and productivity.

### Purpose and Research Questions

The purpose of this study is to empirically compare the informal learning environment preferences of students across creative clusters and STEM clusters in order to generate evidence-based recommendations for discipline-responsive space design. The study is guided by three research questions:

- RQ1: Do students from creative and STEM clusters differ in their demographic composition and patterns of common space use regarding day, time, and purpose?
- RQ2: Do students from creative and STEM clusters differ in their preferences for working style, seating, furniture, and supportive spaces?

RQ3: Do students from creative and STEM clusters differ in their preferences for relaxation behaviour and ambient environmental qualities such as lighting, colour, and atmosphere?

## Theoretical Framework

### Informal Learning Environments and Spatial Theory

Informal learning environments (ILEs) are spaces where collaboration, study, and academic socialisation occur outside scheduled, instructor-directed activities (Jamieson, 2003; Temple, 2008). Moving beyond simple definitions of what the formal classroom is not, contemporary socio-spatial theory defines ILEs by their learner agency. Harrison et al. (2009) conceptualise these settings as "learning landscapes," which are integrated physical and digital ecosystems supporting self-directed, student-led knowledge construction. Unlike a rigid lecture theatre, an informal space offers functional flexibility and variable spatial affordances, fluidly accommodating solo deep-focus work, peer tutoring, and casual interactions (Harrison et al., 2009).

Higher education literature historically treated physical space as a static backdrop, secondary to curriculum or pedagogy (Temple, 2008). Modern scholarship corrects this by drawing on spatial theorists like Henri Lefebvre (1974/1991), who argued that institutional spaces are not neutral containers but are socially produced structures that users actively reappropriate or subvert (Lefebvre, 1992). Applying this concept, Boys (2011) posits that the physical environment operates as a silent pedagogy, communicating implicit messages regarding power, belonging, and institutional validation. Consequently, designing an ILE is a critical pedagogical decision rather than a facilities-management task. As students navigate these fluid spatial boundaries, they actively negotiate their identities as scholars, directly influencing academic engagement and student retention (Boys, 2011; Matthews et al., 2011).

### Disciplinary Cultures and Spatial Implications

Academic disciplines differ not only in what they teach but in how their members think, communicate, and work. Becher's (1989) hard/soft taxonomy captures this distinction: "hard" disciplines (including most engineering and computing fields) tend toward cumulative, quantitative, convergent knowledge-building, whereas "soft" and applied creative disciplines tend toward more iterative, qualitative, divergent, studio-based modes of knowledge-building. Building on this taxonomy, Becher and Trowler's (2001) influential account of "academic tribes and territories" argues that disciplines are not interchangeable containers of content but distinct cultures, each with characteristic epistemologies, communication norms, social structures, and, critically for this study, material and spatial practices. These differing epistemic cultures plausibly translate into differing spatial and material requirements, which the present study tests empirically (Becher, 1989; Becher & Trowler, 2001).

### Reframing Informal Learning Environments as Contemporary Third Places

Informal learning environments can be productively reframed through the "third place," one of the most widely applied contemporary frameworks for understanding informal social space. First articulated in the late 1980s (Oldenburg, 1989, 1999) and since taken up across urban sociology, community planning, retail design, library science, and higher education, the third place describes a category of space distinct from the first place of home and the second place

of work or school. Its core features are consistent: third places are neutral ground entered freely; they act as social levellers, open to all regardless of status; conversation and informal interaction are their primary activity; they keep a low, unpretentious profile; and, cumulatively, they function as a psychological "home away from home."

Viewed through this lens, the informal learning environment is itself a third place within the university campus. Like the cafés the concept originally described, the campus common space sits outside the formal obligations of the classroom yet is not home, and students use it voluntarily for a fluid mix of study, collaboration, rest, and socialising. The qualities that make a third place work, being welcoming, accessible, and low in pressure, are precisely those that make an informal learning environment effective. Reframing the informal learning environment in this way positions it not as leftover study space but as a social setting whose design shapes students' sense of belonging on campus. In the present study, survey items on purpose of use operationalise this third-place function directly, while items on atmosphere and colour tone reflect the concept's emphasis on a playful, low-key, welcoming mood.

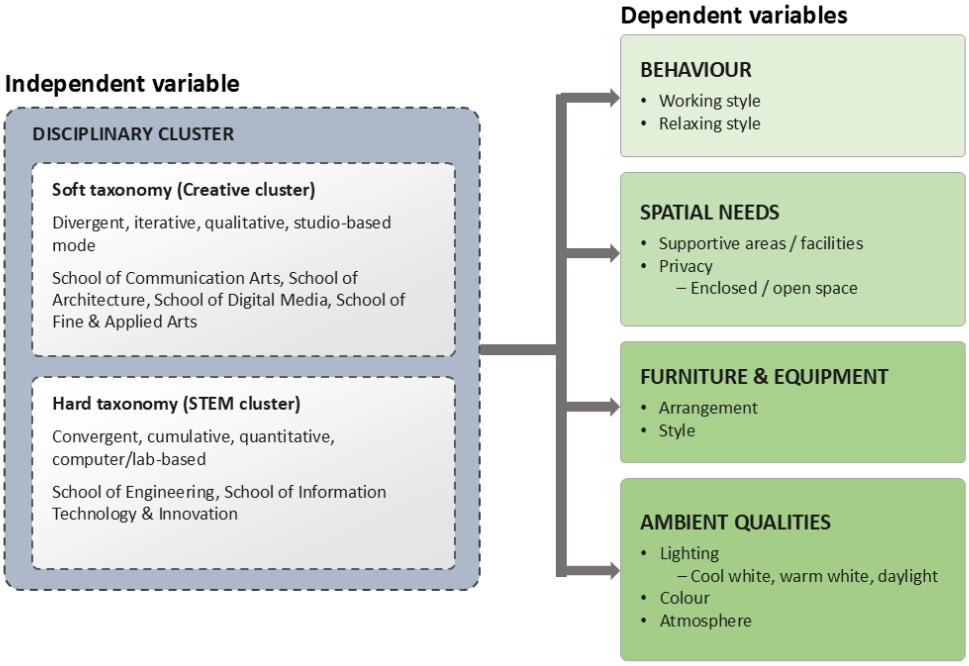
**Environmental Design of Informal Learning Environments**

Informal learning environments can also be understood through environmental psychology, which examines how physical settings shape behaviour, cognition, and wellbeing. A useful organising principle here is person-environment fit: the degree of congruence between an individual's needs and the affordances of their surroundings. From this perspective, a well-designed informal learning environment fits the empirically observed needs of its specific student population rather than a generic "average" user. Evidence from the wider learning-space literature confirms that this fit is consequential: Barrett et al. (2015) demonstrated empirically that physical design features have a significant impact on learning progress, identifying naturalness (including light, air quality, and temperature), individualisation (ownership and flexibility of the space), and stimulation (appropriate levels of colour and visual complexity) as the key contributing dimensions (Barrett et al., 2015).

Within this framework, several interior-design categories are especially relevant to informal learning environments. The first concerns privacy and territorial control. Altman's (1975) theory of privacy regulation conceptualises privacy as a dynamic process of opening and closing the self to others through control over physical boundaries; applied here, a preference for enclosed versus open seating can be understood as a privacy-regulation strategy that may vary with the demands of a student's discipline (Altman, 1975). A second category concerns the provision of distinct functional zones. Prior research consistently reports that students value both quiet, solo-oriented zones for focused work and more open, group-oriented zones for collaboration (Beckers et al., 2016; Harrop & Turpin, 2013), and that spatially well-defined zones support different modes of activity within a single setting. A third category concerns furniture. Flexible, movable, and varied furniture has been associated with greater interaction, collaboration, and engagement, and with students' ability to reconfigure a space to suit the task at hand (Cleveland & Fisher, 2014; Kariippanon et al., 2019). A fourth category concerns atmosphere and ambient conditions, particularly lighting and colour, which have been shown to influence mood, concentration, and comfort in learning settings (Barrett et al., 2015), making them integral rather than incidental to design. Taken together, these categories provide the design vocabulary through which the present study examines whether creative and STEM cluster students differ in the environments that best fit their needs.

Building on the theoretical concepts and empirical findings reviewed above, the key variables and factors were extracted and synthesised into a conceptual framework, with the resulting variable constructs illustrated in Figure 1.

**Figure 1**
*Research Conceptual Framework and Variable Construct*




## Methodology

### Research Design

This study employed a sequential exploratory mixed-methods design comprising a qualitative phase followed by a quantitative phase. In sequential exploratory designs, qualitative data collected in the first phase are used to inform the development of instruments or constructs tested in the second, quantitative phase—a strategy particularly well suited to research questions where the relevant variables are not yet well established in the local context and population.

#### *Phase 1: Qualitative Strand—Instrument Development*

Prior to questionnaire design, in-depth interviews were conducted with students from both the Creative and STEM clusters. The purpose of this qualitative phase was exploratory and instrumental rather than confirmatory: to surface the vocabulary, categories, and behavioural distinctions that students themselves use to describe their informal learning space needs, and to ensure that the subsequent quantitative questionnaire items were grounded in authentic student experience rather than researcher assumptions. The interview data were examined using qualitative analysis, specifically content analysis, to systematically extract the needs, behaviours, and preferences articulated by students within each cluster. The themes generated through this process directly informed the response options used in the multi-select items of the Phase 2 questionnaire (Questions 8, 10, 11, 12, and 19).

### *Phase 2: Quantitative Strand*

**Participants and Sampling.** The quantitative sample comprised 305 students at a private university in Bangkok, recruited via convenience sampling across schools representing two disciplinary clusters. The Creative Cluster (n = 153, 50.2%) comprised Communication Arts (n = 69), Architecture (n = 34), Digital Media and Cinematic Arts (n = 29), and Fine and Applied Arts (n = 20). The STEM Cluster (n = 152, 49.8%) comprised Engineering (n = 96) and Information Technology and Innovation (n = 57). The sample was predominantly Bachelor's-degree students (97.4%), with small numbers of Master's (1.3%) and Doctoral (1.3%) students. Cluster assignment by school produced a near-perfect structural association ($\chi^2(5) = 301.04$, $p < .001$), confirming the cluster-coding scheme functions as a manipulation check rather than a substantive finding.

**Instrument.** The quantitative instrument was a structured, 26-item online questionnaire administered in English. Items 1–5 captured demographic and academic information. Items 6–8 captured patterns of common-space use (day, time, and purpose). Items 9–17 captured work-related preferences (working style, working-pattern behaviours, accessories, supportive space typologies, seating styles, and furniture style and arrangement). Items 18–22 captured relaxation-related preferences. Items 23–26 captured ambient environmental preferences (lighting, colour tone, and atmosphere). Five items (8, 10, 11, 12, and 19) were multi-select; all other items were single-select.

**Data Analysis.** For single-select items, Pearson chi-square ($\chi^2$) tests of independence were conducted on the full contingency table. For multi-select items, each option was tested as a 2×2 table (selected vs. not, by cluster), df = 1. Cramér's V was calculated as effect size throughout ($V \approx .10$ small, .30 medium, .50 large). Where expected cell counts fell below 5 in over 20% of cells, categories were collapsed and re-tested. Given the volume of tests (k = 61), a Bonferroni-corrected threshold ($\alpha = .05/61 \approx .0008$) is reported alongside the conventional $\alpha = .05$ threshold. Analyses were conducted in Python (pandas, SciPy).

## Results

Results are organised by research question: sample composition and usage patterns, work-related preferences, relaxation-related preferences, and ambient environmental preferences. A consolidated summary of all statistically significant and non-significant comparisons is provided in Tables 1 and 2.

### Sample Composition (RQ1)

The two clusters were of comparable size (Creative n = 153, 50.2%; STEM n = 152, 49.8%) but differed significantly in gender composition ($\chi^2(2) = 17.54$, $p < .001$, $V = 0.24$): the Creative Cluster skewed female (62.1%), while the STEM Cluster skewed male. The clusters also differed significantly in year of study ($\chi^2(4) = 14.03$, $p = .007$, $V = 0.22$), with first-year students disproportionately Creative (73.1% of all Year 1 respondents). Programme level did not differ significantly ($\chi^2(1) = 0.12$, $p = .727$, $V = 0.02$).

### Patterns of Common-Space Use (RQ1)

Both clusters showed an almost identical pattern of weekday-dominant use: 88.2% of both groups reported using the common space Monday–Friday ($\chi^2(1) = 0.00$, $p = 1.000$, $V = 0.00$).

Preferred time of use also did not differ significantly ($\chi^2(5) = 7.46$, $p = .189$, $V = 0.16$), with both clusters peaking in the early-to-mid afternoon. Stated purpose of use showed no significant cluster differences across any of the five options tested (all $p > .29$). Both clusters most frequently endorsed "to work" (Creative 63.4%, STEM 59.9%) and "to relax" (Creative 58.8%, STEM 63.8%), consistent with the dual work/leisure function central to Oldenburg's (1989) third-place concept.

## Work-Related Preferences (RQ2)

In contrast to the usage-pattern findings, work-related preferences showed extensive and substantial divergence between clusters.

### *Working Style and Group Size*

Preferred working style differed significantly between clusters ($\chi^2(2) = 12.56$, $p = .002$, $V = 0.20$). STEM students more strongly favoured small-group work (2–4 people: 58.6% versus Creative's 41.2%), while Creative students more strongly favoured large-group work (more than 4 people: 31.4% versus STEM's 15.8%). Preference for solitary, concentration-focused work was similar across clusters (Creative 27.5%, STEM 25.7%).

### *Working-Pattern Behaviours*

Of ten working-pattern behaviours tested, six showed significant cluster differences. STEM students more often reported using shared computers (27.0% vs. 8.5%; $\chi^2(1) = 16.62$, $p < .001$, $V = 0.23$). Creative students more often reported editing video/audio (25.5% vs. 9.9%; $p = .001$, $V = 0.20$), practising speaking/performing (18.3% vs. 6.6%; $p = .003$, $V = 0.17$), drawing/making art (28.1% vs. 13.8%; $p = .003$, $V = 0.17$), shooting content/live-streaming (13.7% vs. 4.6%; $p = .010$, $V = 0.15$), and model-cutting/electronics (19.0% vs. 9.2%; $p = .023$, $V = 0.13$). Behaviours common to both—writing/notes, laptop/iPad use, studying/meetings, brainstorming—showed no differences (all $p > .32$).

### *Accessories and Equipment*

Of eight accessory types tested, three showed significant differences. Creative students were significantly more likely to use cutting mats (19.0% versus 2.6%; $\chi^2(1) = 19.40$, $p < .001$, $V = 0.25$) and content-shooting equipment (24.2% versus 7.2%; $\chi^2(1) = 15.26$, $p < .001$, $V = 0.22$). STEM students were significantly more likely to use power strips (77.6% versus 64.7%; $\chi^2(1) = 5.59$, $p = .018$, $V = 0.14$).

### *Spaces That Support Work*

The strongest disciplinary divergence in the dataset emerged for supportive space typologies: of eight space types, seven showed significant differences. STEM students showed stronger preference for an invention/maker testing area (40.1% vs. 12.4%; $\chi^2(1) = 28.85$, $p < .001$, $V = 0.31$), a brainstorming area (64.5% vs. 45.1%; $p = .001$, $V = 0.19$), and a quick-work area (40.8% vs. 28.1%; $p = .027$, $V = 0.13$). Creative students showed stronger preference for a dedicated relaxation area (13.1% vs. 0.0%; $\chi^2(1) = 19.18$, $p < .001$, $V = 0.25$), an art studio room (38.6% vs. 15.8%; $\chi^2(1) = 18.83$, $p < .001$, $V = 0.25$), a content studio room (35.9% vs. 17.1%; $\chi^2(1) = 12.93$, $p < .001$, $V = 0.21$), and a dressing room (15.0% vs. 4.6%; $\chi^2(1) = 8.21$, $p = .004$, $V = 0.16$). A general activity area showed no difference ($p = .651$).

### *Seating and Furniture*

Seating style for working alone differed significantly ($\chi^2(3) = 8.06$, $p = .045$, $V = 0.16$): Creative students showed stronger preference for enclosed seating (35.9% vs. 27.6%) and standing workspaces (5.9% vs. 1.3%), while STEM students favoured semi-open seating (48.7% vs. 40.5%). Furniture set style also differed significantly ($\chi^2(1) = 5.05$, $p = .025$, $V = 0.13$): both clusters preferred casual over office style, more strongly among Creative students (79.7% vs. 67.8%). Seating for small-group ($p = .661$) and large-group work ($p = .888$), and furniture arrangement (collapsed; $p = .108$), showed no significant differences.

## Relaxation-Related Preferences (RQ3)

Relaxation-related preferences showed almost uniform convergence between clusters. Overall relaxing-style preference did not differ significantly ($\chi^2(2) = 0.85$, $p = .655$, $V = 0.05$), nor did any of the nine relaxation-pattern behaviours (all $p > .09$). Seating style for group relaxation ($p = .311$) and relaxation furniture style ($p = .688$) likewise showed no significant differences. The one significant exception was seating style for relaxing alone ($\chi^2(2) = 8.12$, $p = .017$, $V = 0.16$): Creative students showed stronger preference for fully enclosed seating (62.7% versus 50.0%), while STEM students showed stronger preference for semi-open seating (32.2% versus 18.3%).

## Ambient Environmental Preferences (RQ3)

None of the four ambient-environment items—lighting for the work area ($p = .486$), lighting for the relaxation area ($p = .249$), colour tone ($p = .978$), and overall preferred atmosphere ($p = .331$)—showed statistically significant cluster differences. Both clusters overwhelmingly preferred neutral colour tones (Creative 68.0%, STEM 67.8%) and showed broadly similar distributions across lighting types and atmosphere styles, with “warm and close to nature” and “simple style” most frequently endorsed in both groups.

## Summary of Findings

Table 1 summarises all statistically significant cluster differences ($p < .05$); Table 2 summarises all non-significant comparisons. Of 61 total comparisons, 23 (37.7%) reached conventional significance. Applying the Bonferroni-corrected threshold ($\alpha \approx .0008$) retains nine findings as robust: Gender, School (structural), Working style, Using shared computers, Editing video/audio, Cutting mat, Content-shooting equipment, Invention/maker testing area, and Relaxation area.

**Table 1**
*Statistically Significant Differences Between Clusters (p < .05)*

| Item | Variable / Option | $\chi^2$ | df | p | V | Direction |
|---|---|---|---|---|---|---|
| Q1 | Gender | 17.54 | 2 | < .001 | 0.24 | Creative ↑ female; STEM ↑ male |
| Q3 | Year of study | 14.03 | 4 | .007 | 0.22 | Year 1 ↑ Creative |
| Q5 | School[1] | 301.04 | 5 | < .001 | 0.99 | Structural |
| Q9 | Working style | 12.56 | 2 | .002 | 0.20 | STEM ↑ small groups; Creative ↑ large groups |
| Q10 | Using shared computers | 16.62 | 1 | < .001 | 0.23 | STEM higher |
| Q10 | Editing video/audio | 11.72 | 1 | .001 | 0.20 | Creative higher |
| Q10 | Practising speaking/performing | 8.56 | 1 | .003 | 0.17 | Creative higher |
| Q10 | Drawing/making art | 8.55 | 1 | .003 | 0.17 | Creative higher |
| Q10 | Shooting content/streaming | 6.55 | 1 | .010 | 0.15 | Creative higher |
| Q10 | Model cutting/electronics | 5.20 | 1 | .023 | 0.13 | Creative higher |
| Q11 | Cutting mat | 19.40 | 1 | < .001 | 0.25 | Creative higher |
| Q11 | Content-shooting equipment | 15.26 | 1 | < .001 | 0.22 | Creative higher |
| Q11 | Power strip | 5.59 | 1 | .018 | 0.14 | STEM higher |
| Q12 | Invention/maker testing area | 28.85 | 1 | < .001 | 0.31 | STEM higher |
| Q12 | Relaxation area | 19.18 | 1 | < .001 | 0.25 | Creative higher (STEM = 0%) |
| Q12 | Art studio room | 18.83 | 1 | < .001 | 0.25 | Creative higher |
| Q12 | Content studio room | 12.93 | 1 | < .001 | 0.21 | Creative higher |
| Q12 | Brainstorming area | 10.79 | 1 | .001 | 0.19 | STEM higher |
| Q12 | Dressing room | 8.21 | 1 | .004 | 0.16 | Creative higher |
| Q12 | Quick-work area | 4.89 | 1 | .027 | 0.13 | STEM higher |
| Q13 | Seating — working alone | 8.06 | 3 | .045 | 0.16 | Creative ↑ standing/enclosed |
| Q16 | Furniture style for work | 5.05 | 1 | .025 | 0.13 | Creative ↑ casual style |
| Q20 | Seating — relaxing alone | 8.12 | 2 | .017 | 0.16 | Creative ↑ enclosed; STEM ↑ semi-open |

*Note.* [1] Q5 (School) is a manipulation check; Cluster was operationally defined by School.

**Table 2**
*Non-Significant Differences Between Clusters (p ≥ .05)*

| Item | Variable / Option | $\chi^2$ | df | p | V |
|---|---|---|---|---|---|
| Q2 | Programme (degree level)[2] | 0.12 | 1 | .727 | 0.02 |
| Q6 | Day of common-space use | 0.00 | 1 | 1.000 | 0.00 |
| Q7 | Time of common-space use | 7.46 | 5 | .189 | 0.16 |
| Q8 | To work | 0.27 | 1 | .606 | 0.03 |
| Q8 | To relax | 0.60 | 1 | .437 | 0.04 |
| Q8 | To meet/chat | 0.27 | 1 | .604 | 0.03 |

| | | | | | |
|---|---|---|---|---|---|
| Q8 | Waiting (class/parents) | 1.09 | 1 | .296 | 0.06 |
| Q8 | Practise learning skills | 0.04 | 1 | .845 | 0.01 |
| Q10 | Writing/taking notes | 0.96 | 1 | .327 | 0.06 |
| Q10 | Laptop/iPad use | 0.96 | 1 | .327 | 0.06 |
| Q10 | Studying/online meetings | 0.66 | 1 | .417 | 0.05 |
| Q10 | Brainstorming | 0.03 | 1 | .871 | 0.01 |
| Q11 | Connectors/cables | 1.99 | 1 | .159 | 0.08 |
| Q11 | Writing/flip board | 0.81 | 1 | .369 | 0.05 |
| Q11 | Printer/photocopier | 0.32 | 1 | .569 | 0.03 |
| Q11 | TV/Projector | 0.03 | 1 | .862 | 0.01 |
| Q11 | Headset with microphone | 0.03 | 1 | .869 | 0.01 |
| Q12 | Activity area | 0.20 | 1 | .651 | 0.03 |
| Q14 | Seating — small-group work | 0.83 | 2 | .661 | 0.05 |
| Q15 | Seating — large-group work | 0.24 | 2 | .888 | 0.03 |
| Q17 | Furniture arrangement style[3] | 2.58 | 1 | .108 | 0.09 |
| Q18 | Relaxing style preference | 0.85 | 2 | .655 | 0.05 |
| Q19 | Taking selfies | 2.85 | 1 | .091 | 0.10 |
| Q19 | Watching movies/music/karaoke | 2.77 | 1 | .096 | 0.10 |
| Q19 | Playing games | 2.76 | 1 | .097 | 0.10 |
| Q19 | Snacks/drinks | 1.02 | 1 | .312 | 0.06 |
| Q19 | Chatting with friends | 0.40 | 1 | .527 | 0.04 |
| Q19 | Reading books | 0.20 | 1 | .655 | 0.03 |
| Q19 | Relaxing in quiet atmosphere | 0.16 | 1 | .689 | 0.02 |
| Q19 | Playing music | 0.14 | 1 | .710 | 0.02 |
| Q19 | Napping | 0.03 | 1 | .858 | 0.01 |
| Q21 | Seating — relaxing in group | 2.33 | 2 | .311 | 0.09 |
| Q22 | Furniture style for relaxation | 1.47 | 3 | .688 | 0.07 |
| Q23 | Lighting — work area | 1.44 | 2 | .486 | 0.07 |
| Q24 | Lighting — relaxation area | 2.78 | 2 | .249 | 0.10 |
| Q25 | Colour tone preference | 0.04 | 2 | .978 | 0.01 |
| Q26 | Preferred atmosphere | 5.76 | 5 | .331 | 0.14 |

*Note.* [2] Collapsed categories (Bachelor's vs. Postgraduate). [3] Collapsed categories (duplicate labels merged); uncollapsed test produced a spurious result.

## Discussion

### A Shared Space With Divergent Functional Needs

The results support the conceptual framework. This discussion treats the campus common space as an informal learning environment (ILE) understood through Oldenburg's (1989) "third place," two terms for the same setting viewed from complementary angles: "ILE" names its

educational function, while “third place” names its social role as a space distinct from home and formal obligation. At the level of this broad third-place function, the two clusters are strikingly similar. Both use the ILE predominantly on weekday afternoons, endorse “to work” and “to relax” in near-identical proportions, relax in similar ways, and favour the same neutral palette and warm-natural atmosphere. This convergence is consistent with Oldenburg's (1989) account of a universal need for a neutral, low-pressure, sociable space, one that does not appear to vary by discipline (Oldenburg, 1989).

The clusters diverge sharply, however, in the material and spatial infrastructure required for their discipline-specific work. This divergence is concentrated in the work domain rather than the relaxation domain, and is most pronounced in the “spaces that support your work” construct, where seven of eight options showed significant differences. This selective pattern, convergence in affective and social function alongside divergence in instrumental and material function, refines third-place theory in the higher-education context: the ILE may be simultaneously a unified third place at the social-affective level and a contested, discipline-differentiated territory at the material-functional level.

## Disciplinary Culture in Material Practice

The specific content of the disciplinary divergence aligns closely with Becher and Trowler’s (2001) academic tribes framework. STEM-cluster students’ elevated preference for invention/maker testing areas, brainstorming areas, shared computers, and quick-work areas reflects the structured, cumulative, collectively verified problem-solving characteristic of “hard” disciplines. Creative-cluster students’ elevated preference for art studio rooms, content studio rooms, dressing rooms, and dedicated relaxation areas reflects the embodied, material, and performative knowledge-building characteristic of “soft,” applied creative disciplines—work that requires specialised, often equipment-intensive space that a generic open-plan common area cannot adequately support (Becher & Trowler, 2001).

The enclosure-preference findings add an environmental-psychology dimension to this picture. Creative-cluster students’ consistently stronger preference for fully enclosed seating—evident both in solo work and solo relaxation—may reflect the privacy-regulation demands (Altman, 1975) of creative practice, which often involves vulnerable, in-progress work that benefits from visual and acoustic boundary control. STEM-cluster students’ stronger preference for semi-open seating may instead reflect a working culture oriented toward visibility, peer consultation, and lower-stakes information sharing (Altman, 1975).

## Design Implications

These findings carry direct implications for the planning and design of informal learning environments. First, the strong convergence in ambient and affective preferences suggests that universities can reasonably standardise the base “atmosphere layer” of an ILE (neutral colour tones, natural or daylight-style lighting, and an unpretentious, nature-adjacent or simple aesthetic) across buildings serving different disciplines without risk of alienating either population. Because this layer supports the ILE's shared third-place role rather than any discipline-specific practice, a common baseline is appropriate.

Second, the pronounced divergence in supportive-space typology indicates that the ILE's learning function is not discipline-neutral: a single, undifferentiated open-plan space is unlikely to support the distinct ways Creative and STEM students actually work and learn. Where an

ILE primarily serves Creative-discipline students, planners should prioritise zoned, semi-enclosed or fully enclosed sub-spaces such as a small studio, a content-capture corner, and a quiet relaxation nook. Where it primarily serves STEM-discipline students, planners should prioritise shared technical infrastructure and open brainstorming zones.

Third, for shared or disciplinarily mixed ILEs, these findings argue for a deliberately zoned, hybrid design: a flexible core area satisfying the shared third-place function, surrounded by smaller, differentiated pockets addressing discipline-specific learning needs. Furniture procurement should be diversified accordingly, offering both casual and office-style seating and both enclosed and semi-open configurations in individual-work zones.

### Limitations

Several limitations qualify these findings. First, the study was conducted at a single private university in Thailand, so the specific clusters and patterns reported here may not generalise to institutions with different programme structures or cultural contexts. Second, the sample relied on convenience sampling, and the significant cluster imbalance in the Gender variable may confound findings attributed to discipline; future analyses could incorporate gender as a covariate. Third, because many statistical comparisons were conducted, some individually significant results may reflect chance; the findings emphasised here are those that remained robust across the analysis, while weaker results should be read as suggestive. Fourth, the qualitative phase would benefit from a larger and more varied pool of respondents, allowing a fuller exploration of the range of student lifestyles that shape how these spaces are used. Finally, the cross-sectional design cannot establish whether preferences shift as students progress through their programme.

## Conclusion

This study tested whether Creative and STEM students differ in their preferences for informal learning environments, using a sequential exploratory mixed-methods design in which qualitative interviews informed a 305-respondent survey. The results yield a qualified answer: divergence and convergence coexist, depending on the dimension examined. Demographic composition, working style, working-pattern behaviour, equipment use, supportive-space typology, and solo-seating enclosure each showed significant divergence, most pronounced in the spaces that support work. Ambient and experiential dimensions, by contrast, showed striking convergence: usage timing, broad purpose, relaxation behaviour, and lighting, colour, and atmosphere.

Taken together, these findings suggest the campus informal learning environment is dual in nature. In its social and affective role, it functions as a unified space whose needs converge across disciplines: students of all backgrounds seek the same relaxed, low-pressure setting to work and unwind. In its learning role, it functions as a differentiated space whose needs diverge, because the material and spatial practices through which Creative and STEM students learn are markedly different. A more effective design approach therefore combines a shared, ambience-consistent core serving these universal needs with zoned, discipline-responsive sub-spaces accommodating the specific learning practices of the communities the space serves. Attending to this dual nature, unity in social function and diversity in learning function, offers an evidence-based path toward space that genuinely fits a disciplinarily diverse student body.

## Acknowledgements

The author would like to express sincere gratitude to colleagues Wararat Ponthawee and Nattagan Sangkaew, whose contributions as members of the research team were invaluable to this study. Their assistance, together with that of the interior architecture students who participated in the data collection process, is greatly appreciated.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that AI-assisted writing software was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The author further declares that, apart from Claude, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Altman, I. (1975). *The environment and social behavior: Privacy, personal space, territory, and crowding*. Brooks/Cole.

Barrett, P., Davies, F., Zhang, Y., & Barrett, L. (2015). The impact of classroom design on pupils' learning: Final results of a holistic, multi-level analysis. *Building and Environment*, *89*, 118–133. https://doi.org/10.1016/j.buildenv.2015.02.013

Beagle, D. (1999). Conceptualizing an information commons. *The Journal of Academic Librarianship*, *25*(2), 82–89.

Becher, T. (1989). *Academic tribes and territories: Intellectual enquiry and the cultures of disciplines*. Open University Press.

Becher, T., & Trowler, P. R. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines* (2nd ed.). Open University Press.

Beckers, R., van der Voordt, T., & Dewulf, G. (2016). Learning space preferences of higher education students. *Building and Environment*, *104*, 243–252.

Biglan, A. (1973). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology*, *57*(3), 195–203. https://doi.org/10.1037/h0034701

Boys, J. (2011). *Towards creative learning spaces: Re-thinking the architecture of post-compulsory education*. Routledge.

Cleveland, B., & Fisher, K. (2014). The evaluation of physical learning environments: A critical review of the literature. *Learning Environments Research*, *17*(1), 1–28.

Harrison, A., Duggan, F., Crelin, G., Parekh, N., Austin, S., & Neary, M. (2009). *Learning landscapes in higher education: clearing pathways, making spaces, involving academics in the leadership, governance and management of academic spaces in higher education* [Report]. https://hdl.handle.net/10779/lincoln.24343801.v1

Harrop, D., & Turpin, B. (2013). A study exploring learners' informal learning space behaviors, attitudes, and preferences. *New Review of Academic Librarianship*, *19*(1), 58–77.

Heitsch, E. K., & Holley, R. P. (2011). The information and learning commons: Some reflections. *New Review of Academic Librarianship*, *17*(1), 64–77.

Jamieson, P. (2003). Designing more effective on-campus teaching and learning spaces: a role for academic developers. *International Journal for Academic Development*, *8*(1-2), 119–133. https://doi.org/10.1080/1360144042000277991

Jamieson, P. (2009). The serious matter of informal learning: From the development of learning spaces to a broader understanding of the entire campus as a learning space. *Planning for Higher Education*, *37*(2), 18–25.

Kariippanon, K. E., Cliff, D. P., Lancaster, S. J., Okely, A. D., & Parrish, A.-M. (2019). Flexible learning spaces facilitate interaction, collaboration and behavioural engagement in secondary school. *PLoS ONE*, *14*(10), e0223607. https://doi.org/10.1371/journal.pone.0223607

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall.

Lefebvre, H. (1992). *The Production of Space*. Wiley. https://books.google.co.th/books?id=SIXcnIoa4MwC

Matthews, K. E., Andrews, V., & Adams, P. (2011). Social learning spaces and student engagement. *Higher Education Research & Development*, *30*(2), 105–120.

Oldenburg, R. (1989). *The great good place: Cafes, coffee shops, community centers, beauty parlors, general stores, bars, hangouts, and how they get you through the day*. Paragon House.

Oldenburg, R. (1999). *The great good place: Cafes, coffee shops, bookstores, bars, hair salons, and other hangouts at the heart of a community* (2nd ed.). Marlowe & Company.

Temple, P. (2008). Learning spaces in higher education: An under-researched topic. *London Review of Education*, *6*(3), 229–241.

# AI Literacy, Collaboration, and Future STEM Interests: An Informal Seven-Week Dual-Enrollment Presidential AI Challenge

Elodie Billionniere, Miami Dade College, United States
Lawrence Meyer, Jr., Miami Dade College, United States

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

As artificial intelligence (AI) becomes part of everyday life, preparing young learners to understand and use it responsibly is no longer optional. This study examines the impact of the IDEA-AI: Presidential AI Challenge, an intensive, project-based program designed for dual-enrollment high school students at Miami Dade College. Over a seven-week period, students worked in teams to design AI-informed solutions to real-world community problems. The program integrated foundational AI literacy, ethical reflection (including bias, privacy, and social impact), hands-on prototyping, design thinking workshops, mentoring, and optional industry micro-credentials. Using a structured post-program survey with retrospective pre/post items, we examined changes in students' confidence, understanding of AI concepts, teamwork experiences, and future academic interests. Findings suggest meaningful increases in students' confidence discussing AI, a stronger understanding of how AI can address community challenges, and greater awareness of ethical considerations in AI development. Students reported that collaborative prototyping—including exploratory "vibe coding" sessions—and mentoring strengthened both their technical understanding and communication skills. Importantly, participation also helped many students clarify or deepen their interest in AI, computer science, and related STEM majors. These findings provide insight into how informal, project-based AI programs can responsibly introduce AI education at earlier stages. By intentionally integrating technical learning with ethical awareness and community relevance, such programs demonstrate how AI education can balance technological competence with human-centered thinking. The model offers a practical framework for educators seeking to bridge secondary and higher education while preparing students for thoughtful participation in an AI-driven world.

*Keywords:* artificial intelligence literacy, project-based learning, dual enrollment, informal STEM learning, ethics in AI education, community-based learning

iafor
The International Academic Forum
www.iafor.org

## Introduction

Artificial intelligence (AI) is increasingly shaping communication, healthcare, business, education, and everyday decision-making. As generative AI systems and AI-powered tools become more pervasive, educators and policymakers have emphasized the importance of preparing students to understand, evaluate, and use AI responsibly. Recent scholarship has argued that AI literacy should be considered a foundational competency rather than a specialized technical skill reserved only for advanced computing students (Casal-Otero et al., 2023; Ng et al., 2021).

AI learning experiences are most effective when they combine conceptual understanding with hands-on exploration, ethical reflection, and collaborative problem solving (Leisten & Toth, 2025; Lee et al., 2021). Rather than focusing exclusively on coding or technical implementation, contemporary AI education initiatives emphasize helping students understand how AI systems can be applied to real-world challenges and how AI technologies affect individuals and communities.

Informal and project-based learning environments may be especially valuable for introducing AI concepts to pre-college students. Informal STEM learning environments provide opportunities for experimentation, creativity, collaboration, and mentorship while reducing some of the pressure associated with formal academic assessment. For dual-enrollment students specifically, informal AI experiences may also serve as bridge opportunities that support early exploration of future STEM pathways.

The IDEA-AI: Presidential AI Challenge at Miami Dade College was developed as an informal, seven-week, project-based learning experience for dual-enrollment high school students. The program integrated AI literacy, ethical reflection, design thinking, collaborative problem solving, mentoring, and hands-on prototyping activities centered around community-based challenges. Students worked in teams to develop AI-informed solutions to real-world problems while exploring the social and ethical implications of emerging technologies. The program, and all of its components provided a useful context for examining how informal AI experiences may influence students' perceptions of AI, teamwork, and future STEM interests.

This study examines how participation in the IDEA-AI: Presidential AI Challenge influenced students' AI literacy, collaboration experiences, and future STEM interests. Specifically, the study explores students' self-reported perceptions of learning and development before and after participation in the program. This study was guided by the following research questions:

- **RQ1:** How does participation in an informal, project-based AI program influence students' AI literacy and ethical awareness?
- **RQ2:** In what ways does participation in the program influence students' collaboration and teamwork experiences?
- **RQ3:** How does participation in the program influence students' future STEM interests and academic intentions?

## Program Overview

The IDEA-AI: Presidential AI Challenge was implemented at Miami Dade College as an informal, seven-week learning experience designed for dual-enrollment high school students interested in artificial intelligence, innovation, and community problem solving. The program emphasized applied AI learning rather than advanced technical specialization and was

structured to introduce students to foundational AI concepts in accessible and collaborative ways.

Students participated in team-based activities focused on identifying real-world community challenges and developing AI-informed solutions. Throughout the program, participants engaged in design thinking workshops, collaborative brainstorming activities, mentoring sessions, project proposal development, and hands-on prototyping experiences.

The instructional design intentionally combined technical exploration with ethical reflection. Students examined issues such as fairness, bias, privacy, and social impact while considering how AI systems may influence communities and everyday life. Rather than treating ethics as a separate topic, responsible AI considerations were integrated throughout the project development process.

The program additionally incorporated experiential learning activities using tools such as Google AI Studio, Teachable Machine, Figma, and collaborative design platforms. Exploratory "vibe coding" sessions encouraged experimentation, rapid prototyping, and creative engagement with emerging technologies.

Mentoring was also a central component of the program structure. Students received feedback and guidance from mentors throughout project development, allowing teams to refine ideas, troubleshoot challenges, and strengthen project proposals.

In addition to project-based learning activities, students had opportunities to complete optional industry micro-credentials related to AI and technology. These experiences were intended to support both skill development and future STEM exploration.

The program concluded with collaborative project presentations in which student teams communicated their proposed AI-informed solutions and reflected on the potential community impact of their work.

## Related Work

### AI Literacy in Pre-college and Informal Education

Recent scholarship has increasingly emphasized the importance of AI literacy for broader learner populations beyond computer science majors and technical professionals. Ng et al. (2021) describe AI literacy as a multidimensional construct that includes understanding AI concepts, evaluating AI technologies critically, recognizing ethical implications, and applying AI responsibly within society. Similarly, Whyte et al. (2024) argue that AI literacy initiatives should prioritize accessibility, ethics, and human-centered understanding, particularly for younger learners.

Growing attention has also been given to introducing AI education in K-12 and pre-college contexts. Wishart-Smith (2026) notes that AI education is increasingly moving toward interdisciplinary and socially grounded approaches rather than narrowly technical instruction. This shift reflects concerns that students must be prepared not only to use AI tools, but also to critically examine how AI systems influence communities, workplaces, and everyday decision-making.

Research on AI learning environments further suggests that students benefit from opportunities to engage with AI concepts through hands-on exploration and authentic problem solving (Lee et al., 2021). Rather than teaching AI solely through programming-intensive instruction, many recent initiatives incorporate design thinking, collaborative learning, prototyping, and ethical reflection. These approaches align closely with the structure of the IDEA-AI challenge.

Ethics has also emerged as a central theme in AI education research. Casal-Otero et al. (2023) argue that responsible AI education should include fairness, transparency, accountability, and social impact considerations alongside technical understanding. In the context of this study, ethical AI learning was integrated directly into students' project development process rather than treated as a separate instructional topic.

## Project-Based Learning and Experiential AI Education

Project-based learning continues to be widely recognized as an effective framework for STEM and AI education because it situates learning within authentic and meaningful contexts. Recent work by Kokotsaki et al. (2016) highlights that project-based learning environments can improve student engagement, collaboration, communication, and deeper conceptual understanding when students work toward solving real-world problems.

Project-based learning additionally provides a framework for understanding collaboration within the program. Team communication, shared responsibility, collaborative experimentation, and peer feedback are not simply logistical elements of project work; they are core mechanisms through which learning occurs.

In AI education specifically, project-based and experiential approaches appear especially valuable because AI concepts can be highly abstract for novice learners. Leisten and Toth (2025) found hands-on AI education and first-hand experience at the level of the learner is when AI literacy is most effective. Experiential learning environments may also reduce intimidation associated with emerging technologies by allowing students to explore ideas creatively and iteratively.

The IDEA-AI challenge incorporated project-based learning through team-based problem solving, iterative design, prototyping, presentations, and mentor-supported feedback cycles. Students were asked to identify real-world community challenges and develop AI-informed solutions, reflecting recent recommendations that AI education should emphasize socially relevant applications rather than isolated technical exercises.

## Informal STEM Learning and Interest Development

Recent STEM learning research emphasizes the importance of informal and out-of-school learning environments for supporting student engagement, interest development, and exploration of future academic pathways. Bell et al. (2009) argue that informal STEM experiences can help students connect STEM learning to personal interests, community relevance, and future aspirations.

Informal learning environments may be particularly important in emerging technology fields such as AI due to schools having limited formal AI curriculum opportunities. Short-term enrichment programs, camps, and challenge-based experiences therefore provide opportunities for students to experiment with AI concepts in accessible and low-risk environments.

Research additionally suggests that early exposure to authentic STEM experiences can influence students' future academic interests and confidence. Through engaging students in meaningful problem solving and collaborative learning, informal STEM programs help students better envision themselves participating in future STEM pathways (Dou & Cian, 2021).

In the present study, the IDEA-AI challenge functioned as an informal exploratory experience rather than a formal AI certification or technical training program. As a result, the study focuses on students' self-reported interest development, pathway clarification, and future learning intentions rather than long-term identity formation or persistence outcomes.

### Dual-Enrollment and Transitional Learning Contexts

Dual-enrollment programs occupy a unique space between high school and college. Students are simultaneously developing secondary-level academic identities while encountering postsecondary expectations, disciplinary language, and possible career pathways (Taie & Lewis, 2020). For AI education, this transitional context is important because students may be ready to engage with emerging technologies but may not yet have access to formal AI coursework.

The Presidential AI Challenge therefore functioned as a bridge experience. It introduced dual-enrollment students to AI concepts, collaborative project development, ethical reflection, and future STEM possibilities in an informal setting. The related literature suggests that such bridge experiences can support early exploration and confidence-building, particularly when they combine authentic problems, mentoring, collaboration, and opportunities for students to see the relevance of STEM learning to their own futures.

## Conceptual Framework

This study uses an integrated conceptual framework grounded in three established areas of scholarship: AI literacy, project-based learning, and informal STEM learning. These frameworks were selected because they correspond directly to the design of the IDEA-AI: Presidential AI Challenge and to the constructs measured in the student survey. The framework does not propose a new theory; rather, it applies existing frameworks to an informal, dual-enrollment AI education context.

### AI Literacy Framework

First, Long and Magerko's (2020) AI literacy framework is appropriate because it defines AI literacy as the ability to understand, critically evaluate, communicate about, and use AI technologies in everyday contexts. Their framework emphasizes competencies such as recognizing AI, understanding what AI can and cannot do, interpreting AI outputs, and considering ethical implications.

The AI4K12 framework further supports this study by identifying core ideas that younger learners should understand about AI, including how AI systems perceive, reason, learn, interact, and affect society (Touretzky et al., 2019). These ideas map closely onto the MDC Presidential AI Challenge, which asked students to understand AI concepts, explore AI applications, and consider how AI-informed solutions could address community problems.

In this study, AI literacy is operationalized through survey items that assess students' self-reported understanding of what AI is, confidence discussing AI concepts, understanding of how AI can address real-world problems, awareness of ethical issues, consideration of social impact, and preparedness to use AI responsibly. This operationalization is intentionally broad since the program was designed as an introductory informal experience rather than an advanced technical AI course.

### Project-Based Learning Framework

The second component of the framework is project-based learning. Krajcik and Blumenfeld (2006) describe project-based learning as an instructional approach in which students investigate meaningful problems, collaborate with others, produce artifacts, and develop understanding through sustained inquiry. This framework is applicable because the Presidential AI Challenge was structured around team-based problem identification, design thinking, prototyping, mentoring, and presentation of AI-informed solutions.

Project-based learning serves as the mechanism through which AI literacy was expected to develop. Students did not learn AI only through abstract explanation; they learned by applying AI ideas to community-centered design challenges. The survey items related to their defining of real-world problems, understanding AI components in prototypes, explaining project ideas, developing project specifications, and using prototyping tools are therefore interpreted as indicators of applied project-based learning.

This framework also emphasizes the study's focus on collaboration. In project-based learning, collaboration is not simply a classroom management strategy, it is a central learning process. Students build understanding through communication, shared responsibility, feedback, and iterative refinement. For this reason, survey items related to teamwork, communication, fair distribution of responsibilities, valued contributions, and mentor support are analyzed as part of the project-based learning environment.

### Informal STEM Learning Framework

The third component of the framework is informal STEM learning. The National Research Council's (2009) framework for learning science in informal environments identifies multiple strands of informal science learning, including interest, understanding, participation, reflection, and identity-related outcomes. The National Research Council's later report on productive out-of-school STEM programs emphasizes effective informal STEM programs can support interest, engagement, learning, and future participation across settings (National Research Council, 2015).

The IDEA-AI program was informal, short-term, and exploratory and matches this framework component. The goal was not only to teach AI content, but to also help students explore whether AI and related STEM fields might be relevant to their future academic or career plans. Accordingly, the survey included items on college major interests, academic or career clarification, perceived usefulness of AI for future studies, influence on intended major, and interest in continued AI learning.

The study does not directly measure long-term STEM identity formation or persistence, instead, it examines early indicators of STEM interest development and pathway exploration.

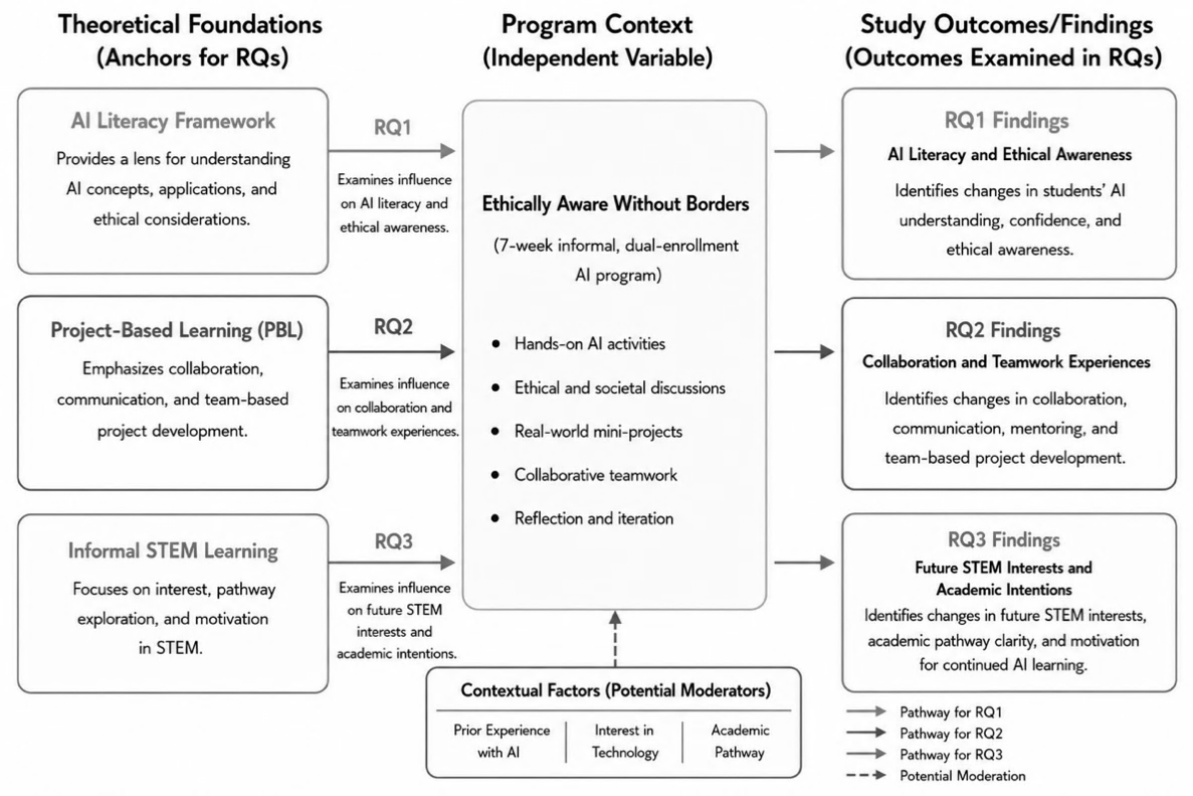
This distinction is important because students' survey responses capture perceived influence immediately following the program, not long-term academic outcomes.

**Integrated Application to the Present Study**

Together, these three frameworks outline the logic of the study. AI literacy identifies the primary learning outcome; project-based learning explains the instructional mechanism; and informal STEM learning explains the relevance of students' future interests and academic intentions.

As shown in Figure 1, the framework aligns directly with the research questions. RQ1 is grounded in AI literacy because it examines changes in AI understanding, confidence, and ethical awareness. RQ2 is grounded in project-based learning because it examines collaboration, communication, mentoring, and team-based project development. RQ3 is grounded in informal STEM learning because it examines future STEM interests, academic pathway clarification, and motivation for continued AI learning.

**Figure 1**
*Integrated Conceptual Framework for the IDEA-AI Study*




This framework is particularly useful for interpreting findings from a dual-enrollment context. Students in the program were not full-time college students but were engaging with postsecondary learning experiences and future academic decision-making. The framework therefore positions the Presidential AI Challenge as a bridge experience supporting early AI literacy, collaborative learning, and STEM pathway exploration.

# Methodology

## Participants and Data Collection

Participants were 18 dual-enrollment students participating in the IDEA-AI: Presidential AI Challenge at Miami Dade College and represented learners at an important transitional stage between secondary and postsecondary education.

Data were collected using a structured post-program survey containing retrospective pre/post Likert-scale items and open-ended response questions. Survey items measured students' perceptions of:

- AI understanding and confidence
- Ethical awareness and responsible AI use
- Collaboration and teamwork
- Project development skills
- Mentoring experiences
- Workshop effectiveness
- STEM interests and future academic intentions

The retrospective pre/post design allowed students to reflect on perceived changes in understanding and confidence resulting from participation in the program.

## Data Analysis

Descriptive analysis was conducted on quantitative survey responses to identify trends across constructs related to AI literacy, collaboration, and STEM interests. Open-ended responses were reviewed to identify recurring themes related to student learning experiences, challenges, and perceived outcomes.

Given the exploratory nature of the study and the relatively small participant group, the analysis focused on identifying patterns and trends rather than establishing causal relationships or statistical generalizability.

# Results

Results are organized according to the three research questions guiding this study. Quantitative findings from retrospective pre/post survey responses are integrated with qualitative findings from open-ended responses to provide a more comprehensive understanding of students' experiences in the IDEA-AI: Presidential AI Challenge.

## Participant Overview

Eighteen students completed the post-program survey. Participants represented students with varying levels of prior familiarity with artificial intelligence, STEM fields, and project-based learning experiences. Survey responses indicated that many students entered the program with interest in technology and innovation, although several participants reported limited confidence in discussing AI concepts prior to participation.

Most participants expressed interest in AI-related or STEM-related majors following participation in the program, including artificial intelligence, computer science, cybersecurity, engineering, and data science.

## RQ1: AI Literacy and Ethical Awareness

The first research question examined how participation in the IDEA-AI program influenced students' AI literacy and ethical awareness. Overall, quantitative findings demonstrated positive shifts across nearly all AI literacy indicators. Students reported increased confidence discussing AI concepts, stronger understanding of how AI can address community challenges, and greater awareness of ethical considerations associated with AI technologies.

### *Understanding Artificial Intelligence*

As shown in Table 1, students reported increased understanding across all AI literacy measures, with the largest gains in confidence discussing AI concepts and understanding of how AI can address community problems.

**Table 1**
*Retrospective Pre/Post Responses on AI Literacy and Understanding*

| Survey Item | Pre: Agree/Strongly Agree | Post: Agree/Strongly Agree |
|---|---|---|
| I understand what artificial intelligence is | 72.2% | 94.4% |
| I understand how AI can address community problems | 66.7% | 94.4% |
| I feel confident discussing AI concepts | 55.6% | 94.4% |
| I feel prepared to use AI responsibly | 61.1% | 88.9% |

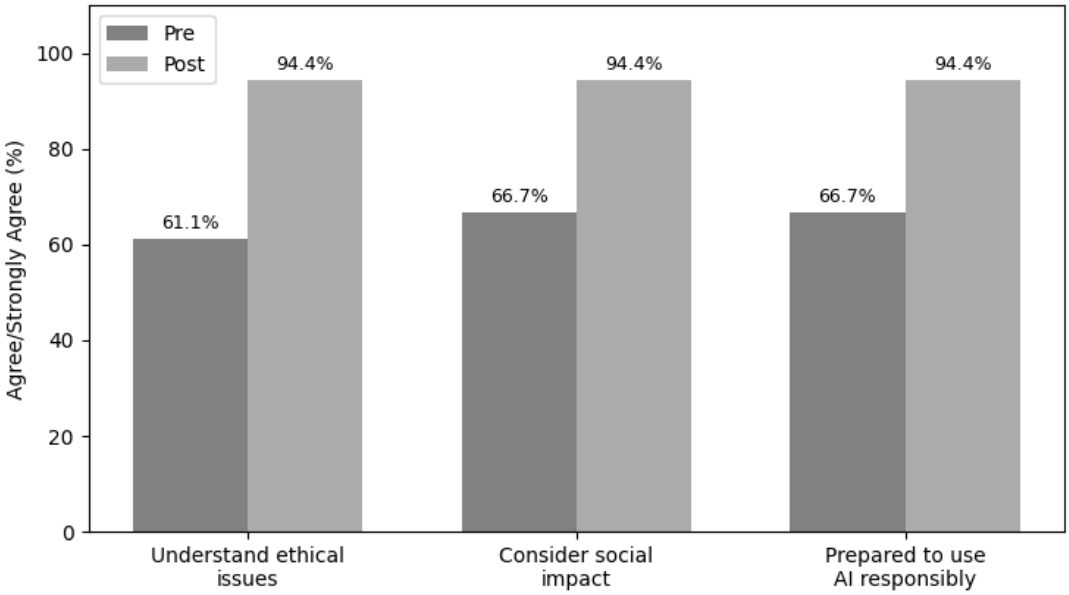
Students' confidence discussing AI concepts demonstrated one of the largest shifts across survey items. Prior to participation, several students selected "Neutral" when asked about confidence discussing AI concepts. Following the program, nearly all students selected either "Agree" or "Strongly Agree."

Similarly, students reported increased understanding of how AI can be used to address real-world community problems. The program's project-based structure appears to have helped students connect AI concepts to socially relevant applications.

### *Ethical Awareness and Responsible AI*

As shown in Figure 2, students reported increases across all AI awareness measures, including fairness, bias, privacy, and social impact.

**Figure 2**
*Retrospective Pre/Post Student Ratings of Ethical AI Awareness*




Across these items, students reported stronger agreement following participation in the program. The largest change appeared in students' understanding of ethical issues related to AI, which increased from 61.1% to 94.4%. Students also demonstrated increased recognition that AI technologies have both technical and societal implications.

Qualitative findings reinforced these trends. Students frequently referenced ethical and community-centered motivations in open-ended responses. One student explained: "AI can be a deeply fulfilling tool to help society if used correctly." Another participant noted: "I learned that AI is not just about technology. It also involves understanding ethics, people, and the impact solutions can have on communities."

Several participants also described entering the program with limited understanding of AI beyond popular media representations. Following the challenge, students described AI as a practical and socially relevant tool connected to real-world problem solving.

One participant reflected: "At first I thought AI was mostly just chatbots and automation. Now I understand how AI can actually help solve community problems when designed responsibly."

These findings suggest that the program supported both conceptual and ethical dimensions of AI literacy.

### *Confidence in Learning Future AI Concepts*

As shown in Table 2, nearly all students expressed confidence in learning more advanced AI concepts and interest in pursuing additional AI learning.

**Table 2**
*Post-program Confidence in Future AI Learning*

| Survey Item | Agree/Strongly Agree |
|---|---|
| I feel capable of learning more advanced AI concepts in the future | 94.4% |
| I am interested in pursuing additional learning related to AI | 94.4% |

These findings suggest that the program may have reduced anxiety associated with AI learning while increasing students' confidence engaging with emerging technologies.

Collectively, the quantitative and qualitative findings related to Research Question 1 indicate that participation in the IDEA-AI challenge contributed to increased AI literacy, stronger ethical awareness, and greater confidence engaging with AI concepts.

## RQ2: Collaboration and Teamwork Experiences

The second research question examined how participation in the program influenced students' collaboration and teamwork experiences. Overall, findings suggest that collaboration was one of the strongest dimensions of the learning experience. Students consistently reported positive experiences related to teamwork, communication, idea sharing, and collaborative problem solving.

### *Teamwork and Communication*

As shown in Table 3, students reported substantial improvements across all collaboration measures.

**Table 3**
*Retrospective Pre/Post Responses on Teamwork and Collaboration*

| Survey Item | Pre: Agree/Strongly Agree | Post: Agree/Strongly Agree |
|---|---|---|
| I work effectively as part of a team | 72.2% | 94.4% |
| I communicate effectively with team members | 66.7% | 94.4% |
| I feel confident explaining project ideas | 72.2% | 100% |
| I can define a real-world problem suitable for an AI solution | 55.6% | 94.4% |

Students additionally reported strong agreement with post-program items related to:

- Shared responsibility among team members
- Feeling that their ideas were valued
- Meaningful contribution to teamwork
- Learning AI concepts more effectively through teamwork

Students also viewed collaboration as important for strengthening communication and project development skills.

### *Mentoring and Collaborative Support*

Mentorship emerged as a significant theme throughout both quantitative and qualitative findings.
Most participants agreed or strongly agreed that:

- Mentoring sessions provided helpful guidance
- Mentors were accessible and supportive
- Mentor feedback improved project quality
- Mentoring helped teams overcome challenges

Qualitative responses further illustrated the importance of mentor support. One participant explained: "Building a project that was applicable to the real world and getting personal feedback from mentors was the most valuable part."

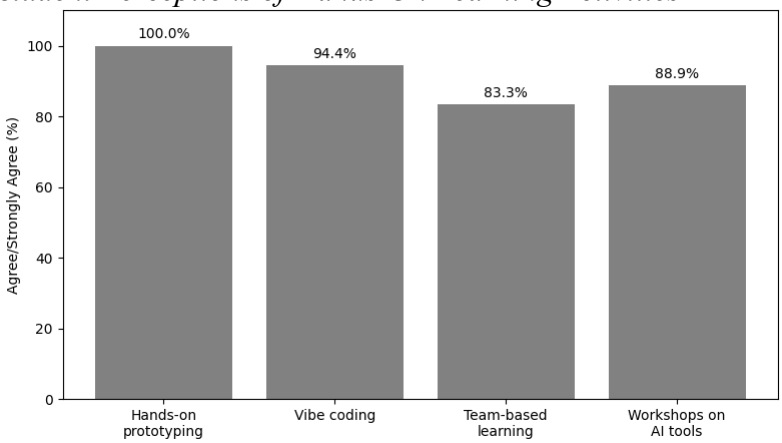
Another student noted: "The mentoring helped us improve our ideas and think more critically about how the solution would actually work."

These findings suggest that mentoring functioned as an important scaffold supporting both collaboration and project refinement.

### *Hands-On Prototyping and Collaborative Learning*

Figure 3 presents students' perceptions of the program's hands-on learning activities, including prototyping and exploratory "vibe coding" sessions, which were reported as highly valuable.

**Figure 3**
*Student Perceptions of Hands-On Learning Activities*




Qualitative responses indicated that students valued opportunities to experiment creatively and collaboratively. One participant stated: "The prototyping activities helped me understand AI better than lectures alone because we were actually building something."

Another reflected: "Working with my team helped me understand different perspectives and improve our project."

These findings suggest that collaborative and experiential learning activities played a central role in supporting student engagement and conceptual understanding.

## RQ3: Future STEM Interests and Academic Intentions

The third research question examined how participation influenced students' future STEM interests and academic intentions. Findings suggest that the program contributed to increased interest in AI and related STEM pathways while also helping students clarify future educational goals.

### *STEM Interests and Academic Pathways*

As shown in Table 4, students reported strong agreement with survey items indicating that participation influenced academic interests and future learning intentions.

**Table 4**
*Student Perceptions of STEM Interests and Future Academic Pathways*

| Survey Item | Agree/Strongly Agree |
|---|---|
| The project helped me clarify my academic or career interests | 88.9% |
| Learning about AI will be useful for my future studies or career | 100% |
| Participation influenced my interest in my intended college major | 83.3% |
| I am interested in pursuing additional AI learning | 94.4% |

Students additionally identified interest in majors and pathways including:

- Artificial Intelligence
- Computer Science
- Cybersecurity
- Data Science
- Engineering
- Information Systems Technology

These findings suggest that the program may have strengthened both awareness of and interest in STEM-related educational opportunities.

### *Community Relevance and Purpose*

An important qualitative theme involved students connecting AI learning to community impact and social relevance. Several students described the value of developing projects connected to real-world needs. One participant explained: "I wanted to work on something that could actually help people in my community."

Another student noted: "The project showed me that AI can be used to solve meaningful problems, not just technical problems."

Students additionally connected AI learning to entrepreneurship, innovation, and future professional opportunities. One participant described having: "A deep passion for building real, tangible products using modern software and artificial intelligence."

Another explained that the challenge strengthened their interest in: "Creating tools to solve real and relevant problems."

These findings suggest that students viewed AI not only as a technical discipline, but also as a socially meaningful and interdisciplinary field.

### *Overall Program Reflections*

Students generally reported positive perceptions of the overall IDEA-AI challenge experience. Most participants agreed that:

- The program met expectations
- They would recommend participation to other students
- Workshops were useful and engaging
- The intensive format supported learning

Students particularly valued:

- Collaborative teamwork
- Mentorship
- Hands-on project development
- Exposure to real-world AI applications
- Opportunities to explore future STEM pathways

Collectively, the findings suggest that participation in the IDEA-AI challenge supported students' evolving STEM interests, future academic aspirations, and understanding of AI as a socially relevant field.

## Discussion

The findings from this study highlight several important considerations for informal AI education in dual-enrollment contexts. Rather than simply demonstrating that students reported positive experiences, the results provide insight into how project-based and ethically grounded AI learning environments may support early engagement with emerging technologies.

### Interpreting AI Literacy Development

The increases observed across AI literacy and ethical awareness measures suggest that introductory AI experiences do not need to rely exclusively on advanced technical instruction to be meaningful (RQ1). Students appeared to develop stronger confidence and broader conceptual understanding when AI learning was connected to authentic community-centered problems.

This finding aligns with recent AI literacy scholarship emphasizing that students should understand AI not only as a computational tool, but also as a social and ethical system that affects communities and decision-making processes (Casal-Otero et al., 2023; Ng et al., 2021). The program's integration of responsible AI discussions throughout project development may have helped students move beyond simplified or media-driven understandings of AI.

Importantly, the findings suggest that ethical reflection can be meaningfully integrated into introductory AI learning environments without requiring separate ethics-only instruction.

Students appeared to interpret ethical awareness as part of the design process itself, particularly when considering fairness, privacy, and community impact.

### Collaboration as a Central Learning Mechanism

The study additionally suggests that collaboration functioned as more than a structural feature of the challenge (RQ2). Team-based learning, mentoring, and collaborative prototyping appear to have shaped how students engaged with AI concepts and project development.

This interpretation is consistent with project-based learning frameworks, which position collaboration, communication, and shared problem solving as central mechanisms for knowledge construction (Kokotsaki et al., 2016). In the context of this study, collaboration may have reduced some of the uncertainty associated with learning unfamiliar AI concepts by creating opportunities for peer discussion, experimentation, and collective exploration.

The role of mentoring is noteworthy. Rather than functioning only as technical support, mentors appeared to support reflection, idea refinement, and problem-solving processes. This suggests that mentorship may be particularly valuable in informal AI learning environments where students are navigating both technical and conceptual uncertainty.

### Informal AI Learning and STEM Pathway Exploration

The findings further suggest that informal AI learning experiences may contribute to students' exploration of future STEM pathways (RQ3). Students' reported interest in continuing AI learning and considering AI-related majors reflects the broader role that informal STEM experiences can play in shaping educational aspirations.

Importantly, this study does not claim a seven-week experience determines long-term STEM persistence or identity formation. However, the findings do suggest that short-duration programs can influence how students perceive the accessibility, relevance, and social purpose of AI-related fields.

The dual-enrollment context may be especially important in this regard. Because students were positioned between secondary and postsecondary education, the program functioned as a bridge experience that exposed participants to collaborative college-level learning environments while allowing exploration of emerging technologies in lower-risk and more exploratory ways.

Students' emphasis on community relevance and socially meaningful problem solving is also significant. Many participants appeared to view AI not solely as a technical discipline, but as a tool connected to helping communities and addressing real-world issues. This finding reinforces recent calls for more human-centered and socially responsive approaches to AI education.

### Implications for AI Education Practice

Several practical implications emerge from this study. First, informal AI programs may serve as effective entry points for introducing students to foundational AI concepts before formal college specialization. Second, combining ethical reflection with project-based learning may help students develop more balanced and socially aware understandings of AI. Third,

collaborative and mentor-supported learning environments may be especially important for novice learners who are encountering AI concepts for the first time. Last, dual-enrollment settings may provide valuable opportunities for connecting secondary students to postsecondary STEM exploration through accessible and interdisciplinary learning experiences.

### Limitations

This study has several limitations. The participant group was relatively small and limited to one institutional context, which restricts generalizability. In addition, the study relied on self-reported survey responses rather than direct assessments of learning outcomes.

The retrospective pre/post design may also introduce response bias, as students reflected on prior understanding after completing the program. Future research could incorporate longitudinal approaches, interviews, or performance-based assessments to better understand how informal AI learning experiences influence student development over time.

## Conclusion

This study examined the impact of the IDEA-AI: Presidential AI Challenge, an informal seven-week AI program designed for dual-enrollment high school students at Miami Dade College. The findings suggest that participation in the program supported growth in AI literacy, ethical awareness, collaboration, and future STEM interests.

Students reported increased confidence discussing AI concepts, stronger understanding of how AI can address real-world challenges, and greater awareness of ethical issues related to AI technologies. Collaborative learning experiences, mentoring, and hands-on prototyping activities were perceived as especially valuable components of the program.

The findings suggest that informal, project-based AI programs may help students clarify future academic interests and explore possible STEM-related pathways. By intentionally integrating technical learning with ethical reflection and community relevance, programs such as the IDEA-AI challenge can support more human-centered approaches to AI education.

As AI continues to influence education and society, there is growing importance in creating accessible learning opportunities that help students engage critically, collaboratively, and responsibly with emerging technologies. Informal dual-enrollment programs may represent one promising model for supporting early AI learning and future STEM exploration.

## Acknowledgements

Partial support for this work was provided by the National Science Foundation Scholarships in Science, Technology, Engineering, and Mathematics (S-STEM) program under Award No. 2220260. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The authors thank the REP Scholars and faculty mentors for their invaluable participation and contributions.

## References

Bell, P., Lewenstein, B., Shouse, A. W., & Feder, M. A. (2009). *Learning science in informal environments: People, places, and pursuits*. National Academies Press.

Casal-Otero, L., Catalá, A., Fernández-Morante, C., Franco-Mariscal, A. J., Fernández-Cerero, J., & Gil-Quintana, J. (2023). AI literacy in K–12: A systematic literature review. *International Journal of STEM Education, 10*, Article 29. https://doi.org/10.1186/s40594-023-00418-7

Dou, R., & Cian, H. (2021). Constructing STEM identity: An expanded structural model for STEM identity research. *Journal of Research in Science Teaching, 59*(3), 458–490.

Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: A review of the literature. *Improving Schools, 19*(3), 267–277. https://doi.org/10.1177/1365480216659733

Krajcik, J., & Blumenfeld, P. (2006). 19. Project-Based Learning. In R. Keith Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 317–333). Cambridge University Press. https://doi.org/10.1017/CBO9780511816833.020

Lee, I., Ali, S., Zhang, H., DiPaola, D., & Breazeal, C. (2021). Developing middle school students' AI literacy. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (pp. 191–197). ACM. https://doi.org/10.1145/3408877.3432513

Leisten, L., & Toth, A. (2025). Towards hands-on learning for supporting AI literacy education across age. *CEUR Workshop Proceedings, 4051*, 60–69. https://doi.org/10.3929/ethz-c-000788637

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). ACM. https://doi.org/10.1145/3313831.3376727

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

Taie, S., & Lewis, L. (2020). *Dual or concurrent enrollment in public schools in the United States* (NCES 2020-125). Institute of Education Sciences. https://ies.ed.gov/use-work/resource-library/report/data-point/dual-or-concurrent-enrollment-public-schools-united-states

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K–12: What should every child know about AI? *Proceedings of the AAAI Conference on Artificial Intelligence, 33*(1), 9795–9799. https://doi.org/10.1609/aaai.v33i01.33019795

Whyte, R., Kirby, D., & Sentance, S. (2024). Secondary students' emerging conceptions of AI: Understanding AI applications, models, engines and implications. In *Proceedings of the 2024 Conference on United Kingdom & Ireland Computing Education Research* (Article 3, pp. 1–7). ACM. https://doi.org/10.1145/3689535.3689552

Wishart-Smith, H. (2026, May 20). *From AI policies to AI literacy in education*. Forbes. https://www.forbes.com/sites/heatherwishartsmith/2026/05/20/from-ai-policies-to-ai-literacy-in-education/

# The Pedagogical Effectiveness of Integrating Gamification and AI Tools in Teaching Physics

Indira Usembayeva, Khoja Akhmet Yassawi International Kazakh-Turkish University, Kazakhstan
Gulzhan Niyazova, Khoja Ahmed Yassawi International Kazakh-Turkish University, Kazakhstan

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

The study is aimed at evaluating the impact of the integration of gamification elements and artificial intelligence tools into the teaching process of the secondary school physics course on students' learning motivation, cognitive engagement, and the quality of mastering the learning material. The main hypothesis of the study is that the systematic use of gamification and AI tools in the educational process contributes to increasing students' interest in the subject and improving their learning outcomes.The research work is based on a quantitative methodology, and the descriptive and cross-sectional research design made it possible to analyze the views and experiences of the research participants within a certain period of time. The obtained results were processed using the Jamovi statistical program and analyzed using a t-test. Along with scientific works published in the last five years, publications from reputable academic databases such as ScienceDirect, Scopus, Web of Science, SpringerLink, and JSTOR were also analyzed. The study describes the scientific and theoretical foundations of gamification technology and analyzes foreign and domestic educational platforms. The work proposes a methodology for creating a digital environment based on gamification technology for teaching physics. The results of the conducted pedagogical research have practical value in creating an innovative educational environment aimed at introducing new teaching methods into the learning process.

*Keywords:* gamification, digital learning environment, game-based learning, motivation, science education

iafor
The International Academic Forum
www.iafor.org

## Introduction

The education sector needs renewal, and countries aspiring to become global economic leaders must begin by building a forward-looking education system. According to the Concept for the Development of Education and Science of the Republic of Kazakhstan for 2023–2029: "The use of digital technologies and artificial intelligence will primarily aim to improve the quality of content, enable personalized learning and assessment tailored to the needs of each student, teacher, and institution, and enhance the educational process" (Government of the Republic of Kazakhstan, 2023).

The ongoing digital transformation in education brings both advantages and challenges. The integration of technology into the sector requires not only technical readiness but also a reconsideration of pedagogical approaches and strategies.

Today, the rapid development and widespread adoption of digital technologies have compelled educational institutions to adapt their teaching methods accordingly. Examples of such adaptation include the use of cloud services, online educational resources, and videoconferencing in the learning process. A core requirement in education is to update learning content while applying effective teaching methods and enhancing their efficiency.

Previous studies have noted that teaching scientific concepts – particularly in science subjects – presents considerable challenges for younger students (Bilal & Erol, 2010). In many countries, including Kazakhstan, science education in general secondary schools relies on traditional textbooks and research materials, which do little to spark students' interest. Although traditional textbooks combine text with visual materials, they often depict static events, and certain scientific concepts cannot be effectively visualized without movement.

For a 7th-grade student who is just beginning to study science subjects, the ability to observe a phenomenon in action and manipulate its parameters increases the interactivity of learning. It has been proven that interactive learning is six times more effective than passive learning (Bayraktar, 2009).

Another challenge in digital education is maintaining students' motivation. For example, global MOOCs (Massive Open Online Courses) provide unrestricted access to educational content at any time and from anywhere in the world. However, research shows that only 7–13% of MOOC users complete their courses, primarily due to a decline in internal motivation during the learning process (Koedinger et al., 2015).

Prior studies indicate that gamification can serve as an effective solution to the issues of motivation and learning efficiency (Ahmed & Asiksoy, 2021; Morales et al., 2014). In education, the term "gamification" refers to the use of game-based technologies in non-game contexts, and scholars consider it a pedagogical technology (Katanosaka et al., 2023). Within the educational context, gamification is widely applied (Zhakupova et al., 2022).

Recent studies demonstrate that pedagogical games enhance students' motivation to learn, as well as their problem-solving skills and creativity. This is largely because games allow students to explore various scenarios and simulate realities that are otherwise impossible to replicate in the real world. Additionally, games enable teachers to monitor student progress. For these reasons, gamification is an invaluable method for allowing students to make mistakes, correct

them, and reinforce learning through diverse scenarios. This is particularly effective for teaching science subjects, which are often perceived as difficult.

Research also indicates that the global gamification market is growing each year. According to Metaari, the use of game technologies in higher education increased by 15.4% between 2019 and 2024 (Ansar & George, 2022). The combination of gamification and group discussion in teaching promotes long-term knowledge retention (Metaari, n.d.).

The core motivational mechanisms of gamification are points, badges, and leaderboards (PBL). Points record a learner's level of engagement and track progress during gameplay. Badges signify a learner's achievements and serve as a primary motivator. Leaderboards display the ranking of individual players or teams (ProProfs Training, n.d.). Gamification is defined as "the use of game mechanics, aesthetics, and game-like thinking to engage people, increase motivation, facilitate learning, and solve problems" (Kononova, 2017). This concept has been applied not only in education but also across various other sectors and educational tools.

Despite the growing research interest and the expansion of practical applications, the concept of gamification remains theoretically underdeveloped, and standardized rules and procedures for its implementation in organizational activities have not yet been established. Gamification is one of the essential functions of any learning management system. In the field of education, challenge-based gamification increases students' academic performance by 34.75%. Students who participate in challenge-based gamified learning improve their performance to 89.45% compared to those who receive only traditional lecture-based instruction (Spinify, n.d.).

Gamification in education, implemented through practical activities, gamified tasks, and group discussions, leads to greater knowledge retention (ProProfs Training, n.d.). Gamification employs active learning principles and spaced repetition, which have been proven to enhance knowledge retention. Gamification technology is being actively integrated into the global education system. Online learning platforms use gamification to attract users and encourage active learning, such as Duolingo, Noon Academy, Kahoot, Quizlet, Memrise, CodeCombat, and EdApp.

In science education, game technologies should be applied not only to simplify complex phenomena but also to help students grasp basic concepts and terms, and to facilitate the memorization of quantities and formulas. Therefore, it is crucial to select games that are aligned with the intended learning objectives.

## Methods and Materials

The purpose of this study is to conduct a comparative analysis of methodologies that incorporate gamification elements in teaching science subjects in domestic and international secondary education institutions. The analysis focuses on the distinctive features of educational models in different countries, teaching formats, the extent of digital resources and interactive tool usage, and aims to identify and present effective practices and innovative solutions. In addition, the study examines the impact of gamification technology on students' motivation and academic performance. The study aims to identify the need for a national gamified digital platform in the Kazakh language designed to enhance learners' motivation and cognitive engagement in natural science subjects, as well as to determine the pedagogical requirements for such a platform.

The conducted study is an exploratory research aimed at determining the scientific and pedagogical justification of the gamified digital platform that the authors plan to develop in the future. That is, the research was carried out as an initial stage that determines what the structure of the future digital platform should be, which functions would be effective to include, and which pedagogical principles and requirements it should comply with.

The main research question of the conducted study is the following: can the pedagogical effectiveness of the digital platform based on gamification technology to be developed in the future be substantiated, and can the existing gamification tools and methods increase the interest of secondary school students in natural science subjects, their activity in class, and their learning outcomes? In addition, the study aims to determine the necessity of a national gamified digital platform in the Kazakh language that increases secondary school students' interest and cognitive activity in natural science subjects, as well as the pedagogical requirements and principles for such a platform.

This research work was carried out as the initial stage of a large-scale scientific and practical project aimed at creating and implementing a gamification-based digital educational platform intended for use in teaching natural science subjects in secondary schools in the future.

The research design is based on quantitative description and cross-sectional reporting. In addition, it is aimed at obtaining objective data about the current potential and pedagogical effectiveness of using gamification technologies in the process of teaching natural science subjects in secondary school.

The methodology of the research is oriented toward obtaining basic information necessary for determining the pedagogical requirements, principles, functional characteristics, and structure of the future gamified and personalized digital educational environment. This approach makes it possible to effectively identify the main patterns in the use of gamification and interactive digital tools in the process of teaching natural science subjects.The platforms ScienceDirect, Scopus, Web of Science, SpringerLink, and JSTOR, among others, were used to search for relevant articles required for the comparative analysis. Various search strategies were applied when investigating methodologies that integrate gamification into the teaching of science subjects in domestic and international secondary education institutions. The main search strategy involved combinations of keywords, which were used to construct advanced searches.

According to statistical data, a 2023 study (Li et al., 2023) conducted a meta-analysis of 41 research papers published between 2010 and 2022 on the use of gamification in education. The results confirmed the effectiveness of gamification in the educational field and identified factors influencing the relationship between gamification and student learning outcomes. Notably, the studies found that gamification proved more effective for middle school students than for high school students, and was particularly successful in teaching science and technical subjects. One of the key works in this area (Sailer & Homner, 2020) carried out an extensive meta-analysis, demonstrating that gamification has the potential to serve as an effective teaching method for activities targeting cognitive, motivational, and behavioral learning outcomes.

Digital simulations are among the most effective gamification tools for mastering educational content. One of the widely used gamification-based systems in science education is PhET Interactive Simulations, developed at the University of Colorado, USA (Wieman et al., 2008). Through interactive simulations in physics, chemistry, biology, and mathematics, this platform

helps students understand complex scientific concepts visually and experientially. It is used in many countries worldwide, including by teachers in Kazakhstan. Another example is Exploration, a "social learning laboratory" well known for its educational projects. By 2024, it had conducted 79 workshops, directly or indirectly reaching 1,591 teachers and over 3 million students (Exploratorium. Impact Report FY24, 2024).

Denmark's Aarhus University developed the nationally recognized ScienceAtHome platform, which includes science-based games and projects (AVID Open Access, 2024). In Japan, the gamified Human Resource Development Game platform, applied in biology and ecology education, is aimed at developing students' logical thinking and systematic decision-making skills. In the UK, MyGCSEScience, adapted for secondary schools, provides video lessons and quizzes in biology, chemistry, and physics, increasing student engagement during lessons (My GCSE Science, 2025). Globally popular resources such as LearningApps.org and Genially are widely used by teachers and students due to their simplicity and the ability to create tasks in the Kazakh language.

At the global level, empirical studies on the use of gamified digital environments in education – particularly in the context of teaching natural science subjects in Central Asian countries – remain limited. Most existing studies focus on the digitalization of education or the use of individual digital tools. Previous research has not considered gamification as a comprehensive pedagogical technology serving as the foundation for designing a dedicated digital educational platform.

Existing educational resources are often insufficiently interactive, especially training simulators that should consist of complex interactive tasks with feedback and analytics. Many current tasks are limited to multiple-choice tests or single-answer questions and demonstrate weak adaptability to students' individual learning potential. As a result, such resources do not adequately support teachers in simplifying routine assessment tasks, nor do they effectively enable learners' self-education, self-directed learning, and self-monitoring. In addition, most of these resources – or at least the most advanced solutions are commercial and require paid access.

The study revealed that most existing educational resources and platforms primarily offer ready-made learning materials and test-based tasks (Meng et al., 2024). However, these approaches fail to ensure full interactivity of the learning process, rapid feedback, and sustained growth in learners' motivation. In this regard, the present study aims to address the identified limitations and to provide an exploratory (diagnostic) rationale prior to the development of a gamified digital educational platform.

Thus, during the theoretical analysis it was determined that at the global level there is a lack of empirical research on the use of gamified digital environments in education, including their application in the context of teaching natural science subjects in Central Asia. The majority of modern studies mainly provide recommendations on the digitalization of education or on the use of individual digital tools. Previous studies have not considered gamification technology as a full-fledged pedagogical technology forming the basis for designing a digital educational platform.

Existing digital interactive educational resources are mostly in the form of simulators, that is, sets of interactive tasks with feedback and analytics (Meng et al., 2024). In most cases, tasks are presented in the form of multiple-choice tests or problems with a single correct answer.

Such tools have a weak potential for increasing students' adaptation and their interest in learning. In such cases, the teacher cannot significantly simplify the daily task of checking the results of students' completed assignments. In addition, most of such resources or advanced solutions are usually paid. Such limitations cannot ensure the full interactivity of the learning process, effective and rapid feedback, and a stable increase in students' interest. Our research complements the limitations identified above and is aimed at creating a preliminary justification for the development of a future gamified digital platform.

In this study, the application of gamification technology in teaching science subjects considered the following requirements: the topic must be based on a learning objective, the expected outcome must be clearly defined, and the activity must correspond to the time allocated for the learning process; the features of gamification technology should be aimed at increasing students' motivation to study the subject and deepening their understanding of the material.

Accordingly, several key features of gamification in teaching science subjects were taken into account:

1. Interactivity of learning materials – Use of simulations and virtual experiments allows students to explore complex phenomena interactively, conduct experiments, and gain a deeper understanding.
2. Group-based gamification – Dividing students into teams to encourage joint decision-making, problem-solving, and practical application of scientific laws, thereby developing communication skills.
3. Increased motivation and introduction of competitive elements – Creating competition among students through gamification to stimulate motivation.
4. Opportunities for research and development – Project-based learning enables students to conduct experiments, apply principles and laws, and analyze results (Tsirulnikov et al., 2023).
5. Development of creativity and thinking skills – Role-playing activities allow students to analyze certain natural processes or unresolved global issues by taking on specific roles, thus fostering creativity.
6. Formation of technical skills – Using gamification technologies and applications improves students' digital literacy and skills in working with IT tools.
7. Visualization of the learning process – Use of infographics, animations, and other visual elements to simplify the explanation of learning material.
8. Assessment and feedback – Interactive quizzes and tests with gamification elements provide instant feedback, making assessment easier and contributing to improved learning quality (Meng, 2024).

In this study, the proposed methods were primarily based on gamification-driven interactive learning applications such as Kahoot, Quizizz, WordWall, and Learning Apps, as well as the integration of riddles and crosswords into the learning process. These applications not only increase student motivation and active participation but are also effective for knowledge consolidation and assessment. The Kahoot platform, for instance, enables teachers to create their own quizzes, present them to students in an engaging format, and assess their knowledge.

The main advantage of the proposed methodology is the consideration of students' individual abilities and characteristics, as well as the use of personalized approaches in the teaching process. When designing the game structure, each student's interests and cognitive abilities are taken into account in advance. This approach encourages independent action, develops critical thinking skills, and consequently increases motivation to master the subject. Thus, the

methodology presented in this study makes the learning process in natural sciences more engaging and effective, while fostering students' subject-specific skills and encouraging active participation.

The research was conducted among 9th-grade students at M. Poshanov Secondary School No. 21 in Turkistan city, Turkistan Region. The study involved two classes from grades 7, 8, and 9 (7A, 7B, 8A, 8B, 9A, and 9B). Specifically, the "A" classes were taught according to the standard curriculum using traditional teaching methods, while the "B" classes used educational games and digital applications in the learning process. The study lasted for one academic quarter, after which the students' level of content mastery, interest in the subject, and class participation were assessed.

In class "A," lessons were conducted in a traditional format: teacher explanations, problem solving, laboratory work, and written assignments. Students listened to the teacher's explanations of new material, took notes using the textbook, and completed exercises at the end of each topic. In class "B," various game-based applications were used in every lesson to complete assignments, review homework, and consolidate new topics.

The practical implementation of the methodology began by integrating gamification elements into traditional teaching. During lessons, students used a specially developed (author's) booklet, which included theoretical materials on the topic "Direct Current," as well as tasks of varying difficulty levels and laboratory work enhanced with gamification elements. The teaching process involved competitive tasks using educational platforms, physics riddles, logical problems, tiered assignments, role-playing activities, group competitions, interactive online quizzes, and quest-based challenges.

As an example, let us consider the topic "Curvilinear motion; uniform circular motion of a material point" taught in Grade 9.

Learning objective: 9.2.1.13 – Describe the uniform circular motion of a body in terms of linear and angular quantities.

For lesson consolidation, a matching task was created in Learning Apps. This activity's advantage is that the teacher can send a link to students for independent review at home, while also monitoring their performance online.

The following lesson was Practical Work No. 1.

In this practical session, before solving problems, students first recalled physical quantities and formulas by playing a game in the Wordwall application. In this activity, students "popped" balloons while recalling the names and symbols of physical quantities. To review formulas, they played the Quiz Show game.

Thus, gamification elements were incorporated into each part of the lesson throughout the learning process.

## Results

In order to assess the level of knowledge acquisition in natural science subjects, including physics, testing was conducted in each of the compared classes. As a result, it was determined

that the classes taught with the active use of gamification elements demonstrated higher results, showed increased interest in the subject, and were more active during lessons. In particular, it was found that the level of knowledge of low-performing students also began to increase.

In addition, a survey was conducted among students in order to determine how traditional teaching and teaching based on gamification technology influence the lesson and to identify the differences in teaching approaches.

The questionnaire developed by the authors consisted of 10 main statements aimed at identifying the effectiveness of using gamification elements in teaching the secondary school physics course, as well as their impact on learning motivation, cognitive activity, and the performance of group work. The responses were evaluated according to the Likert scale: “Strongly agree,” “Agree,” “Disagree,” “Strongly disagree.”

**Figure 1**

*The Impact of Gamification on Learning Motivation and Interest in the Subject*

1.1 1.2

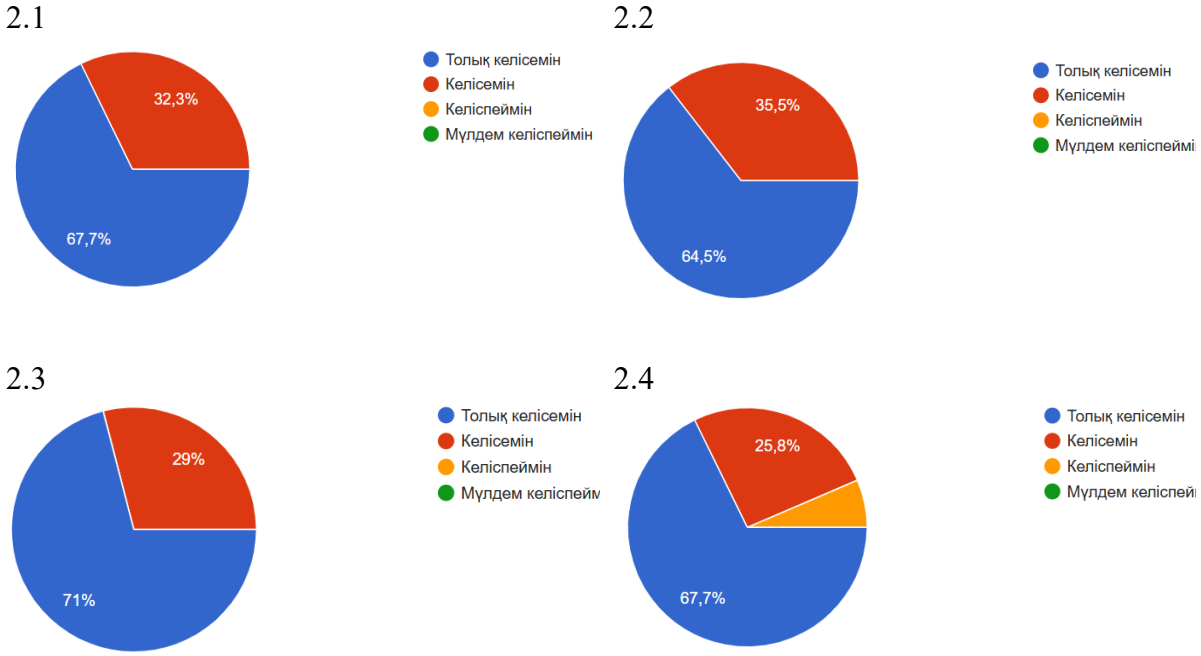
*Note.* 1.1 The use of game elements in physics lessons increases interest in the subject;
1.2 Gamification elements contribute to the formation of a positive attitude toward physics.

The effect of using gamification elements in lessons on learning motivation and interest in the subject is presented in Figure 1. For the two questions, it can be seen that the vast majority of respondents expressed positive opinions. More than 96% of students selected the responses “Agree” and “Strongly agree.”

As shown in Figure 1, the majority of respondents expressed positive opinions regarding the two statements mentioned. In particular, nearly 100% of them marked the responses “Agree” and “Strongly agree.” In this regard, the absence or very small number of negative responses indicates that gamification can be an effective tool for increasing students’ learning motivation.

**Figure 2**

*The Impact of Gamification on Cognitive Engagement and the Learning Process*

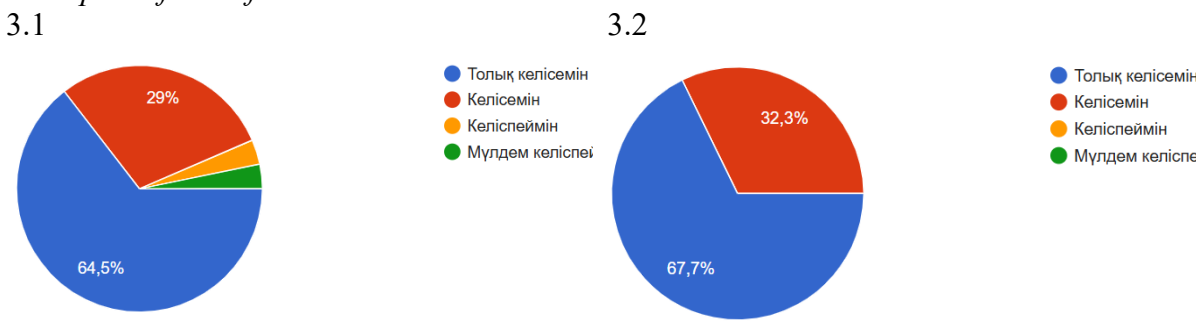



*Note.* 2.1 Gamification techniques such as points, levels, and rewards motivate me to participate more actively in class;
2.2 Gamification makes it easier to understand complex physics concepts;
2.3 Assessing my knowledge is more engaging in lessons that use game elements;
2.4 Gamification helps me maintain attention for a longer period during lessons.

From the diagram in Figure 2, it can be seen that 92–100% of students support the introduction of gamification mechanisms into the learning process and have noticed its positive impact. The highest indicators are related to the fact that using game elements when assessing knowledge is interesting and helps maintain attention for a longer time. The obtained results showed that gamification has a significant role in increasing students' cognitive activity.

**Figure 3**

*The Impact of Gamification on Social and Creative Skills*




*Note.* 3.1 Group competitions and team-based games increase collaboration and active participation during lessons;
3.2 Game-based tasks in physics lessons contribute to the development of my creative abilities.

As shown in Figure 3, according to the two given statements, all respondents 93–100% indicated that gamification elements increase cooperation when working in groups and have a significant impact on their creative abilities. These indicators show that the use of gamification technology in lessons is important not only for developing cognitive skills but also for forming communicative and social skills.

After the completion of the survey, all responses were downloaded and statistical processing of the data was carried out. The statistical analysis was conducted using the specialized Jamovi statistical software. All responses were coded and their percentage indicators were calculated. All empirical indicators were presented visually through the construction of diagrams. The statistical software allows for accurate and clear analysis of the data (Table 1).

**Table 1**
*Statistical Results of Attitudes Toward the Use of Gamification Technology*

| Question N | N | Mean | Mdn | SD | Min | Max |
|---|---|---|---|---|---|---|
| 1 | 62 | 1.36 | 1 | 0.674 | 1 | 3 |
| 2 | 62 | 1.27 | 1 | 0.467 | 1 | 2 |
| 3 | 62 | 1.27 | 1 | 0.467 | 1 | 2 |
| 4 | 62 | 1.64 | 1 | 0.924 | 1 | 4 |
| 5 | 62 | 1.36 | 1 | 0.505 | 1 | 2 |
| 6 | 62 | 1.27 | 1 | 0.467 | 1 | 2 |
| 7 | 62 | 1.45 | 1 | 0.688 | 1 | 3 |
| 8 | 62 | 1.45 | 1 | 0.688 | 1 | 3 |
| 9 | 62 | 1.27 | 1 | 0.467 | 1 | 2 |
| 10 | 62 | 1.45 | 1 | 0.934 | 1 | 4 |

Using the program, descriptive statistical indicators such as the mean value, median, and standard deviation (SD) based on the number of respondents (N) are determined separately for each question. In Table 1, the indicator N shows the number of students who answered a particular question. The mean value reflects the general direction of students' responses. By general direction, we mean the level of students' attitudes toward gamification elements. The median indicates the central value of the statistical data. The standard deviation characterizes the consistency of students' responses.

The results of the statistical data analysis showed that the mean values of all questions ranged between 1.22 and 1.56. The median value was equal to one for all questions, indicating that the majority of students selected the response "Strongly agree." The standard deviation of the statistical data varied between 0.44 and 1.01. This means that although students' opinions were generally in the same direction, there were slight variations in some individual questions.

When calculating the generalized statistical indicators of the survey results, the overall mean value for all questions was 1.37, while the median remained at 1. The standard deviation averaged 0.56. These data indicate that the results obtained during the research are reliable and stable, and that there are no significant deviations among students' responses.

The obtained research results demonstrate that the use of gamification elements in physics lessons increases students' interest in the lesson and their learning motivation. In addition, it

facilitates the understanding of complex topics and concepts, strengthens cognitive activity, and positively influences the development of creative thinking and teamwork skills.

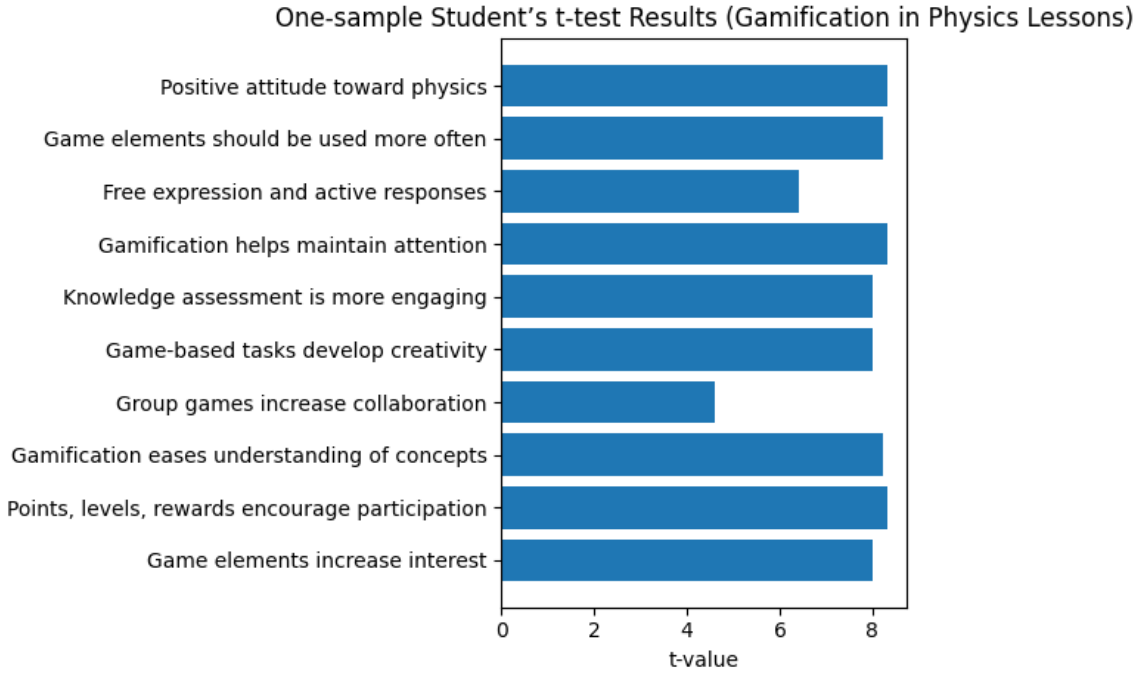
**Table 2**
*One-Sample Student's t-Test, Statistical Results*

| Question | | Statistical data, t | df | p |
|---|---|---|---|---|
| The use of game elements in physics lessons increases interest in the subject. | Student's t-test | 8.00 | 61.00 | < .001 |
| Gamification techniques such as points, levels, and rewards encourage active participation. | Student's t-test | 8.32 | 61.00 | < .001 |
| Gamification makes it easier to understand complex physics concepts. | Student's t-test | 8.22 | 61.00 | < .001 |
| Group competitions and team-based games increase collaboration and engagement during lessons. | Student's t-test | 4.60 | 61.00 | 0.002 |
| Game-based tasks in physics lessons develop creative abilities. | Student's t-test | 8.00 | 61.00 | < .001 |
| Assessing knowledge is more engaging in lessons that use game elements. | Student's t-test | 8.00 | 61.00 | < .001 |
| Gamification helps maintain attention for a longer period during lessons. | Student's t-test | 8.32 | 61.00 | < .001 |
| In such lessons, I freely express my ideas and actively respond to questions. | Student's t-test | 6.42 | 61.00 | < .001 |
| I believe that game and competition elements should be used more frequently in physics lessons. | Student's t-test | 8.22 | 61.00 | < .001 |
| Gamification elements have helped form a positive attitude toward physics | Student's t-test | 8.32 | 61.00 | < .001 |

**Note*. $H_a$ represents the alternative hypothesis, where $\mu \neq 0$.

Using the Jamovi program, a one-sample t-test was conducted to perform a deeper analysis of the data obtained during the study (Table 2). The one-sample t-test makes it possible to determine the significant deviation of the mean value obtained for each question from the accepted null hypothesis (H0: $\mu = 0$). In our research, the alternative hypothesis $\mu \neq 0$ was adopted; therefore, the aim was to determine whether there is a statistically significant difference in the students' responses.

**Figure 4**
*One-Sample Student's t-Test Results*




Let us analyze the obtained results based on the p-value. In Table 2, it can be observed that the p-value is significantly lower than 0.05, which indicates that the responses were not marked randomly and that there is a clear tendency in the students' opinions. For example, in most of the questions the value was $p < .001$, and such a high level of significance shows that students clearly support gamification. All indicators produced statistically significant results. Only one question showed $p = 0.002$, but since it is also lower than 0.05, it is considered statistically significant as well.

The obtained results indicate that the students' responses were not random and demonstrate that students perceive the use of gamification elements in physics lessons as effective. Thus, the results of the statistical test confirm the overall positive results of the survey. Clear evidence was obtained showing that gamification is a factor that increases interest, communication, understanding, sustained attention, and participation in the lesson.

## Discussion

These findings clearly demonstrate both the effectiveness and demand for gamification technology. The survey also confirmed that combining traditional teaching with game elements makes lessons more engaging, more understandable, and helps students absorb information more quickly and effectively.

An analysis of various domestic studies and platforms (such as Online Mektep, Bilimland.kz, BilimCenter.kz, Kitap.kz, Daryn Pro, Roqed, Kundelik.kz, twig-bilim.kz, 100ballov.kz, itest.kz, and others) revealed that, in the process of teaching natural sciences, teachers often need to independently adapt subject content to the individual level of each student, design gamification-based tasks, and effectively integrate them into lessons.

It was observed that both domestic and widely used international educational resources primarily offer ready-made learning materials. The study also found a lack of educational platforms that provide real-time responses to students' actions, deliver immediate feedback, analyze task performance, and offer comprehensive interactive training systems (simulators). Currently, most available tasks are limited to problems and tests with only one correct answer. Such an approach does not fully support the development of students' independent learning skills or allow for a comprehensive assessment of their knowledge.

In addition, many of these platforms are accessible only through paid subscriptions. The methodology for creating a gamified digital environment for teaching natural sciences proposed within this study is aimed at effectively addressing the aforementioned challenges. Future research will continue to explore solutions to this issue.

## Conclusion

This study showed that the development and implementation of a domestic gamification-based digital educational platform can be important for teaching natural science subjects in secondary schools. The research results revealed the high potential of integrating gamification into the educational process through a digital platform. In general, students are ready to use these technologies, but the lack of appropriate platforms and systematic instruction prevents their full implementation in domestic secondary schools.

In conclusion, teaching natural science subjects through gamification technologies can be considered an effective method for ensuring deeper subject understanding, increased student interest, active participation in the learning process, and improved academic performance.

The empirical data obtained during the study made it possible to identify the structure, functional capabilities, and pedagogical requirements of a gamified digital platform to be developed in the future. Conducted as an exploratory study, this research establishes a solid scientific and methodological foundation for the future design and implementation of a gamified digital educational platform for natural science subjects in accordance with the national curriculum.

## Acknowledgement

This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant № AP 26193791).

## References

Ahmed, H. D., & Asiksoy, G. (2021). The effects of gamified flipped learning method on students' innovation skills, self-efficacy towards virtual physics lab course, and perceptions. *Sustainability, 13*, Article 10163. https://doi.org/10.3390/su131810163

Ansar, M., & George, G. (2022). Gamification in education and its impact on student motivation: A critical review. In *Emerging IT/ICT and AI technologies affecting society* (pp. 161–170). Springer.

AVID Open Access. (2024). *ScienceAtHome: Engaging students through science and games*. https://avidopenaccess.org/resource/scienceathome-engaging-students-through-science-and-games/

Bayraktar, S. (2009). Misconceptions of Turkish pre-service teachers about force and motion. *International Journal of Science and Mathematics Education, 7*, 273–291. https://doi.org/10.1007/s10763-007-9120-9

Bilal, E., & Erol, M. (2010). Hypothesis-Experiment-Instruction (HEI) method for investigation and elimination of misconceptions on friction. *Balkan Physics Letters, 18*, 269–276.

Exploratorium. (2024). *Impact Report FY24*. https://www.exploratorium.edu/impact-report-24

Katanosaka, T., Khan, M. F. F., & Sakamura, K. (2023). A physics learning system using gamification for high-school students. In *Proceedings of the 11th International Conference on Information and Education Technology (ICIET)* (pp. 167–171). IEEE. https://doi.org/10.1109/ICIET56899.2023.10111234

Koedinger, K. R., Kim, J., Jia, J. Z., McLaughlin, E. A., & Bier, N. L. (2015). Learning is not a spectator sport: Doing is better than watching for learning from a MOOC. In *Proceedings of the Second ACM Conference on Learning@Scale* (pp. 111–120). ACM. https://doi.org/10.1145/2724660.2724681

Kononova, O. V. (2017). *Proektirovanie informatsionno-obuchayushchey veb-sredy s elementami geymifikatsii. Voprosy organizatsii tekstovogo i igrovogo kontenta* [Designing an information and learning web environment with gamification elements: Issues of organizing textual and game content]. Saint Petersburg. (in Russian)

Li, M., Ma, S., & Shi, Y. (2023). Examining the effectiveness of gamification as a tool promoting teaching and learning in educational settings: A meta-analysis. *Frontiers in Psychology, 14*, Article 1253549. https://doi.org/10.3389/fpsyg.2023.1253549

Meng, C., Zhao, M., Pan, Z., & Bonk, C. J. (2024). Investigating the impact of gamification components on online learners' engagement. *Smart Learning Environments, 11*, Article 47. https://doi.org/10.1186/s40561-024-00336-3

Metaari. (n.d.). *Rynok obuchayushchikh igr vyrastet do 17 mlrd k 2023 godu* [The educational games market will grow to 17 billion by 2023]. https://www.playground.ru/misc/news/metaari_rynok_obuchayuschih_igr_vyrastet_do_17_mlrd_k_2023_godu-311230 (in Russian)

Morales, M., Rizzardini, R. H., & Gutl, C. (2014). Telescope, a MOOCs initiative in Latin America: Infrastructure, best practices, completion and dropout analysis. In *2014 IEEE Frontiers in Education Conference (FIE)* (pp. 1–7). IEEE. https://doi.org/10.1109/FIE.2014.7044096

My GCSE Science. (2025). *About My GCSE Science*. https://www.my-gcsescience.com/

ProProfs Training. (n.d.). *Gamification in training*. https://www.proprofstraining.com/blog/gamification-in-training/

Republic of Kazakhstan. (2023). *Bilim beru zhane gylymdy damytudyn 2023–2029 zhyldarga arnalgan tuzhyrymdamasy* [Concept for the Development of Education and Science of the Republic of Kazakhstan for 2023–2029]. Approved by Resolution No. 774 of the Government of the Republic of Kazakhstan dated September 19, 2023. https://www.gov.kz/memleket/entities/edu/documents/details/505708?lang=kk (in Kazakh)

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review, 32*, 77–112. https://doi.org/10.1007/s10648-019-09498-w

Spinify. (n.d.). *How gamification is different from traditional learning*. https://spinify.com/blog/how-gamification-is-different-from-traditional-learning/

Tsirulnikov, D., Suart, C., Abdullah, R., Vulcu, F., & Mullarkey, C. E. (2023). Game on: Immersive virtual laboratory simulation improves student learning outcomes and motivation. *FEBS Open Bio, 13*(3), 396–407. https://doi.org/10.1002/2211-5463.13567

Wieman, C. E., Adams, W. K., & Perkins, K. K. (2008). PhET: Simulations that enhance learning. *Science, 322*(5902), 682–683. https://doi.org/10.1126/science.1161948

Zhakupova, A., Erkebaeva, S., Karimova, R., & Tezel Shakhin, F. (2022). *Geymifikatsiyanyn bilim beru salasynda koldanu problemalary men damu perspektivalary* [Problems and development prospects of using gamification in education]. *Pedagogy and Psychology, 4*(53), 140–147. (in Kazakh)

**Contact email:** indira.usembayeva@ayu.edu.kz

# Understanding Sustainability in the Context of Anchoring and Objectification: Insights From College Students

Özden Şengül, Boğaziçi University, Türkiye

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Transition to a sustainable society requires transformative changes in people's sense of their role in the natural world rather than a mere technical process. Using anchoring and objectification as lenses, we analyzed student comments in a science education course. The dataset consisted of reflective writings and responses of 163 students from various fields. Using qualitative thematic analysis, this research explores how students consistently anchor sustainability to their lived experiences and enact this approach by translating abstract concepts into concrete behaviors and specific results. Seven major themes were developed as different ways to anchor and objectify the themes. For instance, concepts such as "Waste & Resource Management" link biodiversity, resource conservation, water/energy saving, and waste reduction, as evidenced by observable waste management systems. "Environmental Protection & Biodiversity" situated the preservation of the environment in the context of being a natural outtake of public behavior in shared space with people, invoking clear social norms surrounding what constitutes cleanliness in public space. Conservation of biological diversity was also seen as integral to the structure of education and socialization, as the era when the value was best articulated in "Education & Awareness." Abstract theories of biodiversity become objectified through recycling, waste segregation, and the use of locally sourced products. Lastly, learning nature through sustainability education would be integral to students' ability to construct their understanding of how knowledge is constructed, how values affect research priorities, and how scientific discoveries are told, debated, and re-emerge in public discourse.

*Keywords:* sustainability education, anchoring, objectification, epistemic reasoning, higher education

iafor
The International Academic Forum
www.iafor.org

## Introduction

The 21st century brings challenges that require humanity to aim for a more sustainable society rather than a technical progress (Wals & Jickling, 2002). In this process, people begin to understand their role in the natural world through transformative experiences and imagining and enacting sustainable futures (Sterling, 2011). In college education, students' pre-conceptions about environment and sustainability based on their personal experiences, disciplinary training, social norms, and media (Corcoran & Osano, 2009) influence their understanding and use of sustainability curricula with concrete experiences. There has been less research in higher education regarding how students anchor abstract sustainability concepts into concrete experience, and how they translate or objectify those concepts into observable actions. It is important to understand student cognition to design sustainability education toward informed, evidence-based action.

The present study addresses this gap by examining sustainability awareness among college students through the complementary lenses of anchoring and objectification and epistemic reasoning framework (Kuhn et al., 2000). Through a qualitative study with 163 college students from a large public university in Turkiye, the study demonstrates dominant patterns of sustainability awareness and traces how epistemic reasoning levels shape the ways students connect sustainability to their lives and in their practices. This study was guided by the following questions: What are the dominant patterns of sustainability awareness among college students across thematic categories through anchoring and objectification? How are students' epistemic reasoning levels related to their sustainability awareness?

## Theoretical Framework: Anchoring, Objectification, and Epistemic Reasoning

Anchoring addresses abstract concepts in concrete, lived experience (Hannerz, 2003). Sustainability is an abstract and multi-dimensional construct, and students make sense of it through personal experiences of nature. Anchoring helps students to express what they value about sustainability and capture the interpretive ground from which students reason. In this study, anchoring help use identify the dominant thematic categories about sustainability. It also helps instructors to focus on teaching and learning for professional identity for sustainability. Objectification is also described as a process by which abstract concepts are translated into observable behaviors, practices, and artifacts (Berger & Luckmann, 2016). Objectification is interested in the ways how students render the concept tangible through concrete actions such as recycling, conservation, curricular design, or policy engagement. Objectification is related to practical, action orientation to cultivate sustainability education (Jensen & Schnack, 1997).

This study also employs Kuhn's framework of epistemological reasoning in absolutist, multiplist, and evaluativist levels. This framework aims to define students' epistemic levels to understand how they anchor sustainability concepts and how they objectify them into action. Absolutist level approaches knowledge as objective and certain, multiplist level acknowledges multiple perspectives as equally valid, and evaluativist level supports the critical integration through evidence and argumentation.

## Literature Review

Sustainability education should promote critical thinking and meaningful learning rather than an instrumental approach (Wals & Jickling, 2002). Sterling (2011) defines sustainability education as transformative learning; sustainability education drives learners to reexamine the

core assumptions underpinning their relationship with nature. Critical scholars suggest that mainstream education for sustainable development should move beyond superficial participation and shallower forms of engagement. Kopnina (2020) makes the distinction between rule-following environmental behaviors and deep systemic cognition: learners who lack reflection only implement surface-level practices such as recycling, turning off lights, and using reusable bags.

Students conceptualize sustainability in different ways. For example, Corcoran and Osano (2009) state that students articulate sustainability as a political vision to negotiate tensions between environmental protection, economic development, and social justice rather than a technical matter. The authors also argue that students should be empowered to discuss competing values rather than only internalizing predetermined positions. Although students generally argue for visible, ecological actions such as recycling, they focus less on governance, policy, or economic issues (Acosta Castellanos & Queiruga-Dios, 2022).

Sustainability education also aims to connect knowledge into action. For example, Jensen and Schnack (1997) state that sustainability education should aim to develop students' capacity to make informed, deliberate decisions about sustainability issues and their enactment rather than behavior modification. Therefore, research on transformative learning in environmental contexts suggests tentative and durable changes that can reframe assumptions and integrate new perspectives into identity and practice (Ballantyne & Packer, 2016). Students should justify their choices with evidence rather than counting actions.

Epistemic reasoning is particularly essential for sustainability as a complex and uncertain area. According to Kuhn et al. (2000), individuals may have absolutist conceptions to approach sustainability as a set of rules, multiplist beliefs with multiple equally valid ideas, and evaluativist ideas to evaluate evidence and justify action. This study aims to examine how epistemic levels address the ways students anchor sustainability concepts to lived experience and objectify them into action.

## Methods

A qualitative study was employed during the 2024-2025 academic year. The data collected and analyzed separately and integrated during interpretation. The study was conducted in a science education course at a large public university in Turkiye. The qualitative data collected from 163 students enrolled in Education (ED), Science (SCIE), and Social Sciences (SOC) programs. These students provided reflective writings and open-ended survey responses. Participation was voluntary. Reflective writings invited students to describe how they could contribute to sustainability in their professional roles. Open-ended survey responses elicited students' conceptions of the relationships among science, society, and sustainability. The qualitative materials helped to analyze for anchoring and objectification. Qualitative analysis combined thematic analysis to determine sustainability categories and identify anchoring and objectification patterns. Responses were also coded for epistemic reasoning level using Kuhn et al.'s (2000) framework. Cross-tabulation of epistemic levels with sustainability categories allowed the analysis to trace how reasoning sophistication varied across thematic domains.

## Results

The researchers coded qualitative data based on Kuhn et al.'s (2000) framework. Most students (48.3%) showed an evaluative approach focusing on evidence-based reasoning, critical

evaluation, and integration of diverse perspectives. Some students (33.8%) had multiplist perspectives, recognizing multiple viewpoints without evidence-based evaluation, and less students (17.9%) had absolutist approach and defined sustainability as a set of fixed facts and rules.

Across epistemic levels, "Education and Awareness" theme was the most common (27.6% of coded responses), students from all three departments and all three epistemic levels were interested in teaching as their primary sustainability strategy. On the other hand, Governance, Policy, and Collective Structures had the fewest responses (3.7%) that were predominantly absolutist with lack frameworks for connecting sustainability to governance.

Students with evaluativist perspectives anchored sustainability to systemic, interdisciplinary, and global contexts and objectified it through concrete, innovative actions with an explicit awareness. For example, a student from education major related her future professional identity to biodiversity conservation through curriculum design:

> In the future, as a teacher, I can raise awareness about the importance of ecosystems, genetic diversity, and the interdependence of living organisms. I can contribute to biodiversity education through real-life examples, field trips, and interdisciplinary projects. I believe that I should encouraging critical thinking about human impact on the environment to emphasize students' role in preserving biodiversity. (Education major, female, Education & Awareness theme)

In this quotation, anchoring operated through professional identity as an educator combined with disciplinary knowledge of curriculum design; objectification took the form of specific pedagogical actions, such as curriculum integration, field trips, and critical-thinking exercises, each tied to a reasoned account of how the action produces the intended outcome.

Another example of evaluativist reasoning anchored in disciplinary expertise in trade economics and extended into the governance domain to address a self-critical question about systemic tensions:

> As a future trade expert, I think I can help conserve biodiversity by supporting fair and sustainable trade policies. For example, I would promote eco-friendly products and work with companies that respect environmental standards. Also, I believe international agreements should protect endangered species and reduce overexploitation of natural resources. Trade affects many ecosystems, so we must balance economy and nature. I want to raise awareness in business about how supply chains impact biodiversity. But I still ask myself, can global trade really be sustainable if profit always comes first? (Social Sciences major, male, Governance, Policy & Collective Structures theme)

The student's question presented his evaluativist reasoning since the student proposed actions and interrogated the structural conditions under which those actions would succeed or fail.

Multiplist views referred to the complexity of sustainability and multiple viewpoints although they were equally valid without critically evaluating their validity and effectiveness. In this approach, anchoring was based on personal opinion, classroom learning, and social discourse. For example, in the following reflection, student states:

> I realized that science is not just in labs, it is everywhere in society. I was surprised how much science connects with politics, economy and even ethics. I really enjoyed learning about how science can both help and harm. I still have a question about how we can make science more equal between rich and poor countries. (Education major, male, Conceptual & Theoretical Models theme)

This student explains interdisciplinary connections and concerns about global equity, addressing dual-use approach, help and harm, rich and poor, without an evaluative framework and a concrete plan for action.

Absolutist views defined knowledge as objective, certain, and non-negotiable and presented sustainability as a set of fixed facts and rules. In this view, anchoring defined sustainability as concrete, direct experiences and authority-based knowledge. Objectification included rule-based actions without awareness of complexity. One example illustrated this view:

> I can try to use electricity and water efficiently. I can use reusable packets and bags and containers to decrease amount of plastic usage. I can also be away from giving harm to nature and animals when I am in the outdoors. (Science major, female, Waste & Resource Management category)

This reflection provided objectification of sustainability as rule-following such as delivering information, saving electricity and water, using reusable bags, anchored in basic environmental rules learned through socialization. The actions are unexamined since there is no consideration of effectiveness, or systemic context.

## Discussion and Conclusion

This study showed that students had highest awareness for climate and biodiversity and lowest awareness for governance and individual lifestyle. This indicated that students acknowledged the sustainability as an environmental approach rather than focusing on institutional and economic structures (Acosta Castellanos & Queiruga-Dios, 2022). This approach contradicted with Corcoran and Osano's (2009) explanation that sustainability had political vision for young people since current education needed to cultivate political dimension.

In this study, students anchored sustainability as a pedagogical priority in terms of lived experiences and professional identities (Acosta Castellanos & Queiruga-Dios, 2022). Students focused on informal and formal teaching experiences as a transformative learning opportunity (Sterling, 2011). However, this tendency showed students' overlooking of the policy and economy aspect of sustainability. In terms of objectification, students listed sustainability actions at different epistemic levels. They had multiplist beliefs to recognize the complexity of sustainability without evaluative frameworks (Bendixen & Rule, 2004). Absolutist beliefs aligned with shallower or rule-following understanding. These results suggested that education for sustainable development should target epistemic development alongside content knowledge. As Kopnina (2020) suggested, teachers should teach about sustainability to evaluate sustainability claims, otherwise, it would become rule-following environmentalism.

There were also differences within the departmental level. For example, science students were stronger on environmental categories, social science students were stronger on individual action. Therefore, disciplinary training shaped both content and the sophistication of

sustainability awareness. This finding makes suggestions bridge science and social science perspectives on sustainability to address the government and institutional issues.

This study has implications for future research to track epistemic development over students' academic careers and establish epistemic reasoning shifts. Further research should also compare anchoring patterns across different cultures to understand cultural-specific and cultural-general conceptions. Curriculum design should also focus on preparing governance-oriented sustainability curricula to emphasize policy requirements. This study also has several limitations in terms of single institution sample that limits generalizability to other cultural and institutional contexts. Students in this study enrolled in a science education course: this sample was a convenience sampling within a single course and may introduce self-election bias. In conclusion, these results showed that students who evaluated knowledge claims anchor sustainability more richly and objectify it more responsibly. Sustainability education scaffolded epistemic development and contributed to the transformative change to have transition to a sustainable society.

## References

Acosta Castellanos, P. M., & Queiruga-Dios, A. (2022). From environmental education to education for sustainable development in higher education: a systematic review. *International Journal of Sustainability in Higher Education*, *23*(3), 622–644.

Ballantyne, R., & Packer, J. (2016). Visitors' perceptions of the conservation education role of zoos and aquariums: Implications for the provision of learning experiences. *Visitor Studies*, *19*(2), 193–210.

Bendixen, L. D., & Rule, D. C. (2004). An integrative approach to personal epistemology: A guiding model. *Educational Psychologist*, *39*(1), 69–80.

Berger, P., & Luckmann, T. (2016). The social construction of reality. In W. Longhofer & D. Winchester (Eds.), *Social theory re-wired* (pp. 110–122). Routledge.

Corcoran, P. B., & Osano, P. M. (Eds.). (2009). *Young people, education, and sustainable development: Exploring principles, perspectives, and praxis.* Wageningen Academic Publishers.

Hannerz, U. (2003). Macro-scenarios. Anthropology and the debate over contemporary and future worlds. *Social Anthropology*, *11*(2), 169–187.

Jensen, B. B., & Schnack, K. (1997). The action competence approach in environmental education. *Environmental Education Research*, *3*(2), 163–178.

Kopnina, H. (2020). Education for the future? Critical evaluation of education for sustainable development goals. *The Journal of Environmental Education*, *51*(4), 280–291.

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, *15*(3), 309–328.

Sterling, S. (2011). Transformative learning and sustainability: Sketching the conceptual ground. *Learning and Teaching in Higher Education*, *5*(11), 17–33.

Wals, A. E., & Jickling, B. (2002). “Sustainability” in higher education: From doublethink and newspeak to critical thinking and meaningful learning. *Higher Education Policy*, *15*(2), 121–131.

# How Much Does Financial Support Matter? Changes in Study Abroad Participation and Destination Choice at a Japanese University

Daniel James, Hiroshima Shudo University, Japan

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This paper is the third in a longitudinal series examining the relationship between financial support and study abroad at a medium-sized private university in western Japan. Earlier studies focused primarily on the stated preferences of first-year students before and immediately after the introduction of a revised scholarship system in 2024. The present study shifts the focus from intention to actual participation and compares students who took part in university study abroad programmes in the 2024 and 2025 academic years. The revised scholarship system combines fixed support for first-year students on short-run programmes with variable awards determined by destination, programme duration, and demonstrated foreign-language proficiency or accumulated foreign-language credits. The data show that overall participation increased considerably from 66 students in 2024 to 112 in 2025. The strongest growth occurred in short-run programmes, while exchange participation also increased and expanded to a wider range of destinations. Medium-run programmes remained a small proportion of total participation in both years. Despite these changes in participation, destination preferences were comparatively stable: South Korea remained the leading short-run destination and English-speaking countries continued to attract substantial numbers. The findings suggest that financial support may be more effective in increasing participation than in fundamentally redirecting destination choice. However, the two-year comparison is also likely to reflect exchange-rate conditions, visa requirements, international relations, programme availability, and other institutional and individual factors. Further data from 2026 and 2027 are required to determine whether the observed patterns represent sustained changes.

*Keywords:* Japanese university students, international education, study abroad, scholarships, financial support, student mobility

iafor
The International Academic Forum
www.iafor.org

## Introduction

International education (IE) is an important component of tertiary education in Japan at both national and institutional levels. In 2023, the Government of Japan announced the Japan-Mobility and Internationalisation: Re-engaging and Accelerating Initiative for Future Generations (J-MIRAI), which included a target of 500,000 Japanese students studying abroad by 2033 (Council for the Creation of Future Education, 2023). The target was intended not only to restore outbound mobility after the COVID-19 pandemic but also to strengthen the development of globally capable graduates and Japan's international competitiveness. The scale of the target is ambitious because it requires a substantial expansion of outbound participation in a post-pandemic environment in which tuition, accommodation, travel, insurance, and everyday living costs have risen in many major destinations.

For universities, these national goals are translated into practical questions. Institutions must decide how to encourage students to participate, how to maintain viable partnerships, and how to reduce barriers that may prevent students from studying abroad. Financial support is one of the most direct mechanisms available. Scholarships can reduce the immediate cost borne by students and their families, but they can also be designed as incentives. A scholarship may, for example, reward language achievement, encourage students to study abroad earlier in their degree, make longer programmes more feasible, or support destinations that would otherwise be financially difficult.

The university examined in this study revised its outbound study abroad scholarship system from the 2024 academic year. Before the revision, awards were primarily fixed according to destination and programme duration. From 2024, the system became a hybrid model combining a fixed payment for first-year students participating in short programmes with variable weekly awards linked to destination, programme length, and students' foreign-language proficiency or accumulated foreign-language credits. The rationale was both financial and educational: the university sought to increase outbound participation while also giving students an incentive to improve their language ability and potentially participate in more than one international programme during their undergraduate studies.

The first paper in this series compared first-year student preferences in 2023 and 2024 and suggested that changes in financial support appeared to have less influence on preferred destinations and languages than on timing, programme type, and programme length (James, 2024). The second paper compared first-year cohorts in 2024 and 2025 and again found considerable stability in the popularity of English-speaking countries, South Korea, and China, while noting increased interest in shorter programmes and exchange opportunities. It also highlighted the importance of the weak yen in 2024, when actual outbound participation fell from 84 students in 2023 to 66 in 2024 despite the introduction of the new scholarship system.

The present paper extends this work in an important way. Rather than examining only stated preferences collected through questionnaires, it analyses actual participation in outbound programmes in 2024 and 2025. This distinction is important because students may express interest in several countries or programme types but ultimately participate in only one programme or may decide not to go abroad at all. Actual participation therefore provides a stronger indication of the relationship between institutional financial support and student behaviour, although it cannot establish a simple causal relationship.

The study addresses three questions. First, did overall participation change between 2024 and 2025? Second, did the balance between short-run, medium-run, and exchange programmes change? Third, did the destinations chosen by students change substantially after the revised scholarship system had been in operation for a second year? These questions are considered against the wider context of exchange-rate fluctuations, programme costs, visa requirements, international relations, and the continuing post-pandemic recovery of outbound student mobility.

## Main Body

This is the third paper in a continuing study of the relationship between scholarship design and outbound study abroad at one Japanese university (James, 2024 and 2025). The earlier papers concentrated on the preferences expressed by first-year students and on the institutional reasoning behind the revised scholarship scheme. The 2026 stage of the project moves from preferences to realised behaviour by comparing the number of students who participated in university-administered study abroad programmes during the 2024 and 2025 academic years.

This change of focus allows the scholarship system to be considered from a different perspective. Preference data are valuable because they show what students are considering before detailed programme decisions are made. However, they are necessarily provisional. Actual participation incorporates additional constraints: students must be able to afford the programme, meet academic or language requirements where applicable, obtain a visa when required, fit the programme around their academic schedule, and receive the necessary institutional and family support. The participation data therefore reflect the point at which interest becomes action.

The analysis retains the categories used institutionally. “Seminar” programmes are divided into short-run (SR) programmes of approximately one to ten weeks and medium-run (MR) programmes of approximately three to six months. Exchange programmes normally last six or twelve months. The main comparison is between academic years 2024 and 2025, the first two years in which the revised scholarship system operated.

### Institution in Question and Scholarship Systems: Old, New and Reasoning Behind Change

#### *Institution in Question*

The data were collected at a medium-sized private university in western Japan. Student enrolment is approximately 6,200–6,500 across four undergraduate cohorts. At the time of data collection, the university had 37 partner universities in 17 countries and regions. Its outbound study abroad provision consisted of two principal categories: exchange programmes, generally lasting six months or one academic year, and institutionally organised “seminar” programmes. Seminar programmes were divided into short-run programmes of approximately one to ten weeks and medium-run programmes of approximately three to six months.

Exchange and seminar programmes differ in several important respects. Exchange places are normally limited, often to one or two students per partner institution in an academic year. Applicants must usually satisfy a predetermined language requirement, such as a TOEFL or IELTS score for English-medium study or a TOPIK level for Korean-medium study, and some partners also impose a minimum grade point average. Successful exchange students may take

language courses and, where their proficiency permits, regular academic classes alongside students at the host institution. They continue to pay tuition to their home university and are normally exempt from tuition at the host university.

Seminar programmes generally have lower language-entry requirements and are often focused on language learning, although some also include experiential components such as volunteering, local visits, intercultural activities, or community engagement. Their financial structure is different from exchange. Students remain registered at the home institution and therefore continue to pay home tuition, while programme fees may also include tuition or instructional costs at the host institution. For programmes lasting several months, this creates a substantial additional burden. The difference between exchange and seminar programmes is consequently not simply one of duration. It involves a trade-off between accessibility and cost: seminar programmes are often easier to enter linguistically, while exchange programmes may be financially advantageous once the student has met the academic and language requirements.

These institutional characteristics are central to the interpretation of the 2024–2025 data. A scholarship may reduce the cost of both programme types, but it interacts differently with each. For a short-run seminar, a fixed award can reduce the immediate price of a relatively short overseas experience. For an exchange programme, a performance-related award may reinforce an existing financial advantage created by the tuition-waiver arrangement. For medium-run seminars, by contrast, scholarship support may still be insufficient to offset the combined effect of programme fees, home tuition, accommodation, travel, and living costs.

### *Scholarship Systems: Old, New and Reasoning Behind Change*

Prior to 2024, the university operated a non-repayable scholarship system in which the amount awarded was largely determined by programme destination and duration. Exchange students received fixed amounts for six- or twelve-month participation, while seminar students received fixed payments according to the length and geographical area of the programme. The system was straightforward, but it did not directly reward students for improving their language proficiency and was increasingly challenged by rising programme costs.

The financial environment changed considerably after the COVID-19 pandemic. Tuition, accommodation, insurance, and travel costs increased in several of the university's established destinations, particularly English-speaking countries. At the same time, the weak yen increased the Japanese-yen cost of overseas study. These conditions risked making study abroad less accessible, especially for students considering programmes of several months or more. The university therefore redesigned its scholarship system with two related objectives: to increase outbound participation and to connect financial support more closely to students' academic and language development.

From 2024, the new system used a staged model. Weekly scholarship amounts varied according to destination and the student's stage, with the stage determined through recognised foreign-language test results or the number of foreign-language credits earned at the university. Higher stages attracted larger weekly awards. The system was intended to make language development materially meaningful: improving a test score or accumulating language credits could increase the scholarship available for a later study abroad programme.

The university also recognised a potential disadvantage for first-year students. New entrants have had less time to accumulate university language credits and may not yet have obtained

high scores on external proficiency tests. To avoid discouraging early participation, first-year students on programmes of three months or less were made eligible for a one-time fixed award of ¥50,000, unless their stage-based entitlement produced a higher amount. This element was designed to make short-run participation more accessible at an early point in the undergraduate degree.

The revised system also allowed students to receive support for more than one overseas programme during their university studies. This was intended to create a developmental pathway: a student might begin with a short-run programme, return with greater motivation, improve language proficiency, move to a higher scholarship stage, and later apply for an exchange or another programme. In this sense, the scholarship was designed not only as financial assistance but also as an incentive structure.

The earlier studies found that the immediate effect of the new system was difficult to isolate. Actual participation fell from 84 students in 2023 to 66 in 2024, a decline that was interpreted primarily in relation to the weak yen and rising overseas costs rather than as evidence that the scholarship reform had failed. At the same time, questionnaire data suggested growing interest in short programmes and exchange programmes, while destination preferences remained comparatively stable. The 2025 participation data therefore provide the first opportunity to examine whether a second year of the revised system coincided with a substantial change in realised participation.

## Methodology

The present study uses descriptive institutional data on students who participated in outbound study abroad programmes in the 2024 and 2025 academic years. Unlike the 2024 and 2025 papers, which analysed questionnaires completed by first-year students, this study does not measure hypothetical or stated preferences. It compares actual participation by programme category and destination.

The principal variables were academic year, programme type, programme duration category, and destination. Programmes were classified according to the university's existing categories: exchange, short-run seminar, and medium-run seminar which were described above. The analysis is descriptive. Counts and percentage changes are used to identify differences between the two academic years, and destination distributions are compared to assess whether increased participation was accompanied by a substantial redistribution of students across countries.

The total number of outbound participants was 66 in 2024 and 112 in 2025. In 2024, 3 students participated in exchange programmes and 63 in seminar programmes. In 2025, 8 participated in exchange programmes and 104 in seminars. Within the seminar category, there were 59 short-run and 4 medium-run participants in 2024, compared with 97 short-run and seven medium-run participants in 2025.

When given the prompt, "Summarize the data for participation into a table form suitable for an academic article written on APA7 guidelines" ChatGPT produced Tables 1 and 2 that were later redrawn by the author for formatting purposes (Open AI, 2026).

**Table 1**
*Study Abroad Participation by Programme Type*

| Academic Year | Exchange | Seminars | Total |
|---|---|---|---|
| 2024 | 3 | 63 | 66 |
| 2025 | 8 | 104 | 112 |

**Table 2**
*Seminar Participation by Duration*

| Academic year | Short-run | Medium-run | Total seminars |
|---|---|---|---|
| 2024 | 59 | 4 | 63 |
| 2025 | 97 | 7 | 104 |

Several limitations should be made clear. First, the study concerns one private university and should not be treated as representative of all Japanese higher education. Second, only two years of actual participation under the revised scholarship system are compared. Third, the data are observational and do not permit a causal claim that scholarship reform produced the changes. Exchange rates, programme availability, visa policies, international relations, student finances, family circumstances, language confidence, academic scheduling, and employment concerns may all influence participation. Fourth, programme destinations are partly constrained by the university's partner network and the programmes offered each year. The findings should therefore be interpreted as institutional trends that can inform further longitudinal research rather than as a controlled evaluation of scholarship effectiveness.

## Results

Overall participation increased from 66 students in 2024 to 112 in 2025, a rise of approximately 69.7%. This brought outbound numbers closer to the institution's post-pandemic recovery trajectory. The increase occurred in both seminar and exchange participation, although the numerical growth was dominated by short-run seminars.

Short-run seminar participation rose from 59 students in 2024 to 97 in 2025, an increase of approximately 64.4%. Medium-run participation increased from four to seven students but remained a small component of overall outbound mobility. Medium-run students represented approximately 6.1% of all outbound participants in 2024 and 6.3% in 2025. Thus, despite a small increase in the raw number, the proportional position of medium-run study changed very little and remained below 10% in both years.

Exchange participation increased from three students in 2024 to eight in 2025. More importantly for the institutional analysis, the range of exchange destinations widened. The 2024 exchange group was distributed across South Korea, New Zealand, and Canada, with one student in each destination. In 2025, South Korea accounted for five exchange students, while New Zealand, Australia, and the Czech Republic each received one. Australia and the Czech Republic therefore appeared as additional exchange destinations in the 2025 data.

Destination patterns for seminar programmes showed both growth and continuity. South Korea remained the largest single short/medium-run destination, increasing from 19 students in 2024 to 36 in 2025. New Zealand also remained highly attractive, rising from 23 to 31. Participation in the United States increased from eight to 16 and the United Kingdom from two to nine. France changed from four students to three, while Spain increased from one to five and

Germany from one to two. China was the principal exception to the general growth pattern, declining from five students in 2024 to two in 2025.

The destination data therefore show that the large rise in participation was not accompanied by a wholesale shift toward a new geographical pattern. The same broad destination groups continued to be most popular; nearby South Korea and established English-speaking destinations. The clearest change was the increased number of students within those existing patterns, together with some diversification in exchange destinations.

When given the prompt, "Summarize the data for destination into a table form suitable for an academic article written on APA7 guidelines" ChatGPT produced Tables 3 and 4 that were later redrawn by the author for formatting purposes. (Open AI, 2026)

**Table 3**
*Seminar Destinations*

| Destination | 2024 | 2025 |
|---|---|---|
| South Korea | 19 | 36 |
| China | 5 | 2 |
| New Zealand | 23 | 31 |
| United States | 8 | 16 |
| United Kingdom | 2 | 9 |
| France | 4 | 3 |
| Spain | 1 | 5 |
| Germany | 1 | 2 |

**Table 4**
*Exchange Destinations*

| Destination | 2024 | 2025 |
|---|---|---|
| South Korea | 1 | 5 |
| New Zealand | 1 | 1 |
| Australia | 0 | 1 |
| Canada | 1 | 0 |
| Czech Republic | 0 | 1 |

## Discussion

The most important finding is the scale of the increase in actual study abroad participation. The number of students rose from 66 in 2024 to 112 in 2025, an increase of approximately 70%. This is significant in the context of the earlier papers because 2024 had initially appeared disappointing: participation fell even though the revised scholarship system had just been introduced. The 2025 result suggests that the first year should not be treated as a definitive test of the scholarship reform. Instead, it supports the earlier argument that the effects of financial incentives need to be examined longitudinally and in relation to external economic conditions.

One plausible explanation for the contrast between the two years is the exchange-rate environment. James (2025) identified the weak yen as a major factor behind the decline from 84 participants in 2023 to 66 in 2024. A scholarship denominated in yen can become less effective when overseas tuition, accommodation, and living costs rise sharply in yen terms. Conversely, when students perceive the exchange-rate environment as less adverse, the same

scholarship structure may have greater practical value. The data in Figures 3 and 4 shows the 2025 increase as occurring after the particularly difficult financial conditions of 2024. This does not demonstrate that exchange-rate movement caused the increase, but it reinforces the need to understand institutional scholarships in relation to the wider cost of international mobility.

The second major finding concerns programme duration. Short-run participation increased from 59 to 97 students, approximately 64%. This was the largest numerical contribution to the overall rise. The result is consistent with the design of the revised scholarship system, particularly the fixed ¥50,000 award available to first-year students on programmes of three months or less. Short-run programmes already have structural advantages: they require less time away from Japan, can often be scheduled during university vacations, involve lower total accommodation and living costs, and are less likely to conflict with academic progression, club commitments, or employment preparation. A fixed scholarship can therefore remove or reduce a marginal financial barrier for students who are already interested but uncertain about the cost.

This pattern also supports a distinction between encouraging participation and transforming the type of international experience students choose. The scholarship system may be most effective where the student is relatively close to deciding to participate. For a two- to ten-week programme, a ¥50,000 payment can represent a marked reduction in the total cost. For a programme lasting several months, the same absolute amount would be a much smaller proportion of the financial commitment. Performance-related weekly awards may provide more support for longer participation, but students still face home tuition, host costs, accommodation, travel, and daily living expenses. Consequently, the relative effect of financial support is likely to vary according to programme duration.

Medium-run programmes remained comparatively unpopular. There were four participants in 2024 and seven in 2025, meaning that medium-run programmes accounted for less than 10% of all outbound participation in both years. This stability is noteworthy because the revised system was intended to provide greater support according to programme length and student achievement. The data suggest that financial support alone may not overcome the structural disadvantages of the medium-run category. Such programmes are long enough to generate substantial tuition, accommodation, and living costs but do not necessarily provide the tuition-waiver advantage associated with exchange. They may also be difficult to fit around the Japanese academic calendar. In some destinations, visa requirements create an additional barrier. The United States is a particularly relevant example because longer periods of study can involve more complex visa procedures than short visits. Medium-run programmes can therefore occupy an awkward middle position: substantially more expensive and administratively demanding than short-run study, but without the financial and academic benefits of formal exchange.

Exchange programmes also grew. Between 2024 and 2025 the increase was approximately 267%, while the actual number rose from three students in 2024 to eight in 2025. The direction of change is clear: more students participated in exchange in 2025, and the range of destinations expanded to include Australia and the Czech Republic. This is important because exchange programmes have relatively high entry barriers. Students must normally reach specified language levels and may compete for a small number of places. However, once accepted, they generally pay tuition only to the home university. The revised scholarship system may strengthen the attractiveness of this route by rewarding the same language achievement that

students need to qualify for exchange. In other words, the incentive and the entry requirement point in the same direction.

An equally interesting finding is what did not change. Despite the revised scholarship system and the substantial increase in total participation, destination preferences remained remarkably stable. South Korea continued to be the most popular destination for short-run programmes. This is understandable in practical terms: it is geographically close to Japan, travel times and transport costs are relatively low, and the Korean language and culture have strong popular appeal among Japanese university students. These characteristics existed before the scholarship revision and are not fundamentally altered by the amount or structure of financial support.

English-speaking countries also remained consistently attractive. New Zealand, the United States, and the United Kingdom all received substantial numbers of seminar participants in 2025, while Australia appeared as an exchange destination. This continuity mirrors the preference data in the earlier papers, in which the traditional English-speaking destinations remained popular even when students were aware, or would later become aware, of their comparatively high costs. English occupies a distinctive position in Japanese education, and study in an English-speaking environment carries linguistic, academic, and symbolic value. A scholarship can make these destinations more financially accessible, but the underlying attraction is likely to be driven by factors extending beyond cost.

This stability has an important implication for scholarship policy. Financial incentives appear capable of changing whether some students participate without necessarily changing where they most want to go. In economic terms, the scholarship may affect the threshold decision to study abroad more strongly than the ranking of destinations. Students may use additional support to realise an existing preference rather than to select a destination solely because it offers a larger scholarship. This interpretation is consistent with the earlier questionnaire studies, which repeatedly found that destination preferences were more stable than preferences concerning timing, programme length, and programme type.

China was the main exception. Seminar participation declined from five students in 2024 to two in 2025 and diplomatic tensions may have contributed to this decrease. Such an interpretation is plausible, but it should remain cautious because the numbers are small and the study does not include interview data explaining individual decisions. Perceptions of safety, bilateral relations, media coverage, programme timing, cost, and alternative destinations may all affect participation. The key point is that a financial explanation alone is insufficient. A destination may become more or less attractive because of political or social conditions even when scholarship support remains available.

The findings therefore reinforce the distinction developed across this series of studies. Scholarship design matters, but it operates within a larger decision-making environment. Porter and Porter (2020) identify non-financial influences on Japanese students' study abroad decisions, including language confidence, university club responsibilities, employment concerns, and reluctance to spend extended periods away from family. The present data are compatible with that wider view. Short-run programmes can reduce several of these non-financial barriers simultaneously: they require less time away, interfere less with recruitment and academic schedules, and demand a smaller total financial commitment. The scholarship may strengthen these advantages, but it does not create them.

The institutional implications are therefore twofold. First, if the immediate objective is to increase the number of students gaining some form of overseas experience, targeted support for short-run programmes appears promising. The 2025 participation increase suggests that a relatively accessible programme combined with visible financial support can attract substantially more students. This can be valuable for students who would not initially consider a semester or year abroad and may provide a first step toward later international engagement.

Second, if the objective is to increase participation in medium-run or exchange programmes, financial support should be combined with other measures. For exchange, language preparation, clear proficiency targets, academic advising, and early promotion may help students reach the entry threshold. For medium-run programmes, institutions may need to examine the total cost structure, credit recognition, academic-calendar fit, and visa support rather than if a larger scholarship alone will make the category attractive. The persistently small medium-run share suggests that the programme model itself deserves review.

Nevertheless, the two-year period is too short for strong causal conclusions. The 2024 cohort experienced an unusually weak yen, while the 2025 environment was perceived as financially less severe. Programme availability may have changed, new partner relationships may have opened opportunities, and visa policies can alter the feasibility of destinations. International relations may also affect perceptions of certain countries. The observed increase should therefore be described as occurring under the revised scholarship system, not as being caused solely by it.

Further longitudinal analysis is essential. Data from 2026 and 2027 will make it possible to determine whether the high level of short-run participation persists, whether exchange numbers continue to grow, and whether medium-run programmes remain below 10% of total mobility. A four-year dataset will also reduce the extent to which a single exchange-rate shock or a small number of individual destination choices dominate the interpretation. In addition, post-return surveys or interviews should ask students how important the scholarship was in their final decision, whether they would have participated without it, and which other factors influenced programme and destination choice. Similar data from students who considered study abroad but did not participate would be particularly valuable.

Overall, the 2024-2025 comparison suggests a scholarship system that is associated with greater participation but comparatively stable destination choice. That pattern is consistent with the policy objective of widening access to study abroad, but it also demonstrates the limits of financial incentives. Money matters, particularly in lowering the barrier to participation, but it operates alongside academic, linguistic, cultural, political, and personal considerations.

## Conclusion

This study compared actual outbound study abroad participation at one medium-sized private Japanese university during the first two years of a revised scholarship system. The clearest finding was the substantial rise in participation from 66 students in 2024 to 112 in 2025. Short-run programmes accounted for most of this growth, increasing by approximately 64%, while exchange participation also increased and included a wider range of destinations. Medium-run programmes, by contrast, remained relatively unpopular and represented less than 10% of total participation in both years.

The findings suggest that financial support appears to encourage more students to participate in study abroad, particularly in shorter programmes where the overall financial commitment and other barriers are lower. The fixed award for first-year students may be especially relevant because it provides a clear and immediate reduction in cost at an early stage of university study. The performance-related elements of the scholarship system may also support exchange participation by rewarding the language achievement that students need to meet exchange-entry requirements.

However, financial support alone does not appear to fundamentally change where students want to study. South Korea remained the dominant short-run destination, and English-speaking countries continued to attract substantial numbers. The decline in participation in China, despite the broader increase elsewhere, further suggests that destination choice can be shaped by diplomatic, social, and perceptual factors that cannot be explained by scholarship amounts alone.

The study remains limited by its two-year timeframe, single-institution setting, and descriptive design. Exchange rates, visa policies, international relations, programme availability, and individual student circumstances may all have influenced decisions. The particularly weak yen in 2024 is an important example of an external factor that may have reduced the practical value of institutional financial support.

For these reasons, data collection will continue in 2026 and 2027. A longer longitudinal dataset will help determine whether the growth in short-run and exchange participation represents a sustained trend or a short-term response to changing economic and international conditions. Future research should also combine participation statistics with interviews or questionnaires for both participants and non-participants. Such evidence would allow a more precise assessment of when financial support changes a student's decision, when it simply helps realise an existing preference, and when non-financial barriers remain more influential. For institutions seeking to expand outbound mobility, the present findings indicate that scholarships are an important tool, but they are most effective when understood as one component of a broader system of academic, linguistic, administrative, and personal support.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Chat GPT, a Large Language Model, was used in production of the tables used in the manuscript. The usage was limited to producing the data for the tables, however, the formatting i.e. the APA style as is laid down by the American Psychological Association (APA) Publication Manual of the American Psychological Association, Seventh Edition, was carried out by the author. The author further declares that, apart from Chat GPT, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Council for the Creation of Future Education. (2023, April 27). *Japan-Mobility and Internationalisation: Re-engaging and Accelerating Initiative for future generations (J-MIRAI) outline*. Cabinet Secretariat.

James, D. (2024). Does financial support matter? Scholarships and study abroad preferences for first year university students in Japan. *The European Conference on Education 2024: Official Conference Proceedings,* 80. https://doi.org/10.22492/issn.2188-1162.2024.80

James, D. (2025). Study Abroad and Scholarships: Are Student Preferences Affected by Performance Related and Determined Financial Incentives? *The European Conference on Education 2024: Official Conference Proceedings*, 673. https://doi.org:10.22492/issn.2188-1162.2025.53

OpenAI. (2026). *ChatGPT* (August 8 Version) [Large Language Model]. https://chatgpt.com/chat

Porter, P., & Porter, N. (2020). Japanese college students' study abroad decisions: Perspectives of Japanese study abroad administrators. *International Education Journal: Comparative Perspectives, 19*(2).

# Does AI Ideation Feel Collaborative? Collaboration Potential as a Structurally Distinct Evaluation Dimension in an LLM-Based Ideation Tool

Stefania Zourlidou, University of Koblenz, Germany
Kamyab Farokhi, University of Koblenz, Germany
Frank Hopfgartner, University of Koblenz, Germany

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Large language models are increasingly used to help students generate project ideas, but an idea that is easy to generate is not necessarily one that students see as useful for working with others. This paper re-analyses a pilot dataset from EduIDEAtor, an LLM-based ideation tool, to ask a narrow question that was not addressed in two earlier papers from the same deployment: does perceived collaboration potential move together with other quality ratings, or does it show a different response pattern? Sixteen students completed the post-use survey, with 14 valid responses for collaboration potential. Collaboration was rated positively: 78.6% selected Better or Much Better than their usual brainstorming method (median = 4 on the intended five-point scale). However, collaboration potential was weakly related to comparative engagement (rho = .27), creativity (rho = -.02), and ease of use (rho = .20). Its only clear association within the comparative block was with project relevance (rho = .65, p = .013). It was not significantly associated with any of the seven broader experience items. Reliability of the five-item comparative block increased from alpha = .814 to .860 when collaboration was removed. With this small sample and a single collaboration item, the result does not establish a separate construct. It does suggest that perceived collaboration potential may capture something different from general usability and engagement, and should be tested directly in future group-based studies.

*Keywords:* large language models, collaboration, data science education, ideation, project-based learning

iafor
The International Academic Forum
www.iafor.org

## Introduction

Large language models are increasingly discussed as tools for teaching and learning. They can support tasks such as generating educational content, answering questions, and providing tutoring support, but they also create risks around inaccurate information and misuse. Kasneci et al. (2023) discuss opportunities, challenges, risks, and mitigation strategies for LLMs in education. Lo (2023) similarly describes uses such as instructor assistance and virtual tutoring, while noting the risk of incorrect or fabricated information. Both reviews discuss learner- and instructor-facing uses of LLMs. The present paper turns to a narrower issue: how students judge the potential use of AI-generated ideas with other people. Project-based learning is designed to engage students in the investigation of authentic problems (Blumenfeld et al., 1991). In collaborative learning, however, putting learners in the same situation does not guarantee that productive collaboration will occur. Dillenbourg (1999) argues that collaborative situations are expected to trigger particular forms of interaction, but these interactions may or may not actually happen. For an ideation tool, this means that generating a useful idea for one student is not the same as supporting discussion and joint development of that idea. Creativity-support research makes a similar distinction. Shneiderman (2007) describes creativity support tools as useful for individual expression, group collaboration, and social creativity. Frich et al. (2019) provide a broad review of 143 HCI papers on creativity support tools and map key characteristics of this design space. Combined, these results suggest that creativity support is not limited to a single mode of use. For an LLM ideation tool, this raises the practical question of whether a student may like the interface and the ideas it produces while still judging those ideas differently when considering their use in a group setting.

### Novel Contributions Relative to Earlier Publications

This paper uses the same EduIDEAtor pilot dataset (Farokhi & Zourlidou, 2026) as two earlier papers, but it does not repeat their main analyses. Farokhi et al. (2026) reported the broad pilot evaluation: usability and experience ratings, comparisons with familiar brainstorming, and associations with satisfaction and willingness to reuse the tool. That paper already showed that collaboration potential received positive ratings in the comparative block. A second paper from the same deployment focused on open-ended improvement comments and translated them into interface and implementation priorities, such as recoverability, clearer output presentation, transparency, and more steerable literature suggestions (Zourlidou et al., 2026). The present paper focuses on a more specific question. It does not examine whether students liked EduIDEAtor, whether they would use it again, or which interface improvements should be prioritised. Instead, it concentrates on a narrower aspect of their experience: does perceived collaboration potential rise and fall with engagement, creativity, ease of use, and project relevance, or does it behave differently? The earlier pilot paper reported the overall reliability of the five comparative items, but it did not examine their inter-item structure or what happens to reliability when the collaboration item is removed. The design-oriented paper did not address this question either.The present paper therefore examines whether perceived collaboration potential moves in line with the other comparative quality ratings or instead shows a relatively distinct response pattern within this pilot dataset.

## Method

The analysis uses 16 anonymised post-use survey responses collected after students interacted with EduIDEAtor at a German university. Full recruitment, procedure, and tool details are reported in Farokhi et al. (2026), and the anonymised dataset is publicly available (Farokhi &

Zourlidou, 2026). In brief, EduIDEAtor supported early project ideation by taking a student's interests and constraints and returning a small set of candidate directions with short explanations and suggested first steps. The evaluated version was used in a single session before the survey was completed. The survey contained five comparative questions asking students to rate EduIDEAtor against their usual brainstorming method on engagement, creativity, ease of use, project relevance, and collaboration potential. The response anchors were Much Worse, Worse, Same, Better, and Much Better. For this analysis, the verbal responses were coded directly using the released codebook: 1 = Much Worse, 2 = Worse, 3 = Same, 4 = Better, and 5 = Much Better. Coding from the original labels avoids ambiguity about the neutral midpoint and preserves the ordinal meaning of the questionnaire. One respondent left the complete comparative block blank. A second response recorded collaboration potential as the invalid combined value “Same;Better”; this entry was treated as missing for analyses involving collaboration. Therefore, collaboration analyses use n = 14. The survey also contained seven five-point experience items covering navigation, input clarity, satisfaction, engagement, match to interests and goals, creativity support, and motivation.

The analysis was deliberately small and descriptive. Spearman rank correlations were used because the survey responses are ordinal and the sample is small. First, collaboration potential was correlated with the four other comparative dimensions. Second, it was correlated with the seven broader experience items. To show the contrast, we also inspected the six pairwise correlations among engagement, creativity, ease of use, and project relevance. Finally, Cronbach's alpha was calculated for the complete five-item comparative block and again after removing collaboration potential. Reliability calculations use the 14 respondents with complete data on all five comparative items. The p-values are reported as descriptive aids rather than as confirmatory evidence. The sample is too small for strong inference, and several related correlations are inspected. The interpretation therefore focuses on the size and pattern of associations, not on a pass/fail significance rule.

## Results

### Collaboration Ratings

As already suggested by the broad pilot evaluation, perceived collaboration potential was not a weakly rated item. Among the 14 valid responses, the median was 4.0 (IQR = 0.75), and 11 students (78.6%) selected Better or Much Better than their usual brainstorming method. The new result is not this positive average. It is the way the collaboration ratings relate to the rest of the survey. Collaboration potential had weak associations with comparative engagement (rho = .27, p = .357), creativity (rho = -.02, p = .956), and ease of use (rho = .20, p = .492). Project relevance was the exception: it showed a moderate-to-strong positive association with collaboration potential (rho = .65, p = .013). The seven broader experience items showed the same general picture. None had a statistically significant association with collaboration potential in this small sample, although satisfaction (rho = .47) and motivation (rho = .44) showed the largest positive coefficients.

**Table 1**
*Spearman Associations With Perceived Collaboration Potential*

| Variable | n | rho | p |
|---|---|---|---|
| Comparative: Engagement | 14 | .267 | .357 |
| Comparative: Creativity | 14 | -.016 | .956 |
| Comparative: Ease of use | 14 | .201 | .492 |
| Comparative: Project relevance | 14 | .646 | .013 |
| Experience: Easy to navigate | 14 | .169 | .563 |
| Experience: Input was clear | 14 | .166 | .570 |
| Experience: Satisfaction | 14 | .469 | .090 |
| Experience: More engaging than traditional | 14 | -.140 | .633 |
| Experience: Ideas matched interests/goals | 14 | .348 | .223 |
| Experience: Creativity support | 14 | .337 | .239 |
| Experience: Motivation | 14 | .439 | .117 |

*Note.* Pairwise available cases are used. The p-values are descriptive and are not adjusted for multiple comparisons.

## Non-collaboration Ratings

The four non-collaboration comparative dimensions showed a clearer shared pattern. Their six pairwise Spearman correlations ranged from rho = .47 to .80. Five of the six were statistically significant at $p < .05$: engagement-creativity (.80), engagement-ease of use (.79), creativity-ease of use (.73), creativity-project relevance (.54), and ease of use-project relevance (.58). The remaining association, engagement-project relevance, was positive but weaker (rho = .47, p = .088). In other words, the non-collaboration ratings generally rose and fell together, while collaboration potential joined that pattern only through project relevance.

## Reliability Check

The reliability check points in the same direction, but should be interpreted cautiously. The five comparative items had Cronbach's alpha = .814. When collaboration potential was removed, alpha increased to .860. Removing an item does not prove that the item measures a separate construct, especially with only 14 complete cases. Here it is simply another indication that collaboration potential adds a different response pattern to an otherwise more internally consistent group of ratings.

**Table 2**
*Reliability of the Comparative Evaluation Block*

| Scale | Items | Cronbach's alpha |
|---|---|---|
| Full comparative block | Engagement, creativity, ease of use, project relevance, collaboration potential | .814 |
| Without collaboration potential | Engagement, creativity, ease of use, project relevance | .860 |

*Note.* Both estimates use the same 14 complete comparative responses.

## Discussion

The main result is that students could rate collaboration potential positively without rating it in the same way as ease of use, creativity, or engagement. Those other quality judgments mostly moved together. Collaboration did not follow that pattern as consistently. Its strongest and only statistically clear link within the comparative block was project relevance. One possible interpretation is that students use different information when they judge whether an AI-generated idea could work in a group. Interface quality may matter for the experience of generating an idea, while project relevance may matter more when deciding whether the idea is worth taking to peers. This interpretation is compatible with Dillenbourg's (1999) point that collaboration depends on interactions that actually occur, not simply on placing learners in a collaborative situation. It is also consistent with Shneiderman's (2007) distinction between individual expression, group collaboration, and social creativity. This interpretation nevertheless must remain tentative. The survey did not ask students to explain what they meant by collaboration potential, and it did not observe any actual group interaction. The item may reflect imagined shareability, perceived usefulness for discussion, or something else. The data show a response pattern, but they do not tell us the psychological reason for that pattern.

The result suggests a useful design hypothesis. Improving navigation, wording, or visual polish may improve the individual experience without necessarily changing perceived collaboration potential. If future versions of EduIDEAtor aim to support group work, collaboration should be designed and evaluated directly rather than treated as a side effect of good usability. Possible features include a simple way to share an idea card, compare two students' candidate ideas side by side, add short peer comments, or export selected ideas as discussion prompts. These are proposals for future testing, not features shown by the present study to be effective. Their value would need to be assessed through actual collaborative tasks. This is also where the present paper differs from the earlier design-oriented analysis. Zourlidou et al. (2026) used student comments to identify general interface priorities such as recoverability, clearer presentation, transparency, and steerable literature suggestions. The current analysis does not replace those priorities. It adds a separate question for the next design cycle: if collaboration matters, what specific interaction makes an AI-generated idea easier to discuss, compare, revise, or adopt with others? A similar point applies to classroom use. An individual AI ideation step can be followed by a deliberately social step: for example, students can bring one or two candidate ideas to a short peer-review activity and explain why the idea is relevant, feasible, and worth developing. This kind of activity would make collaboration observable rather than assumed. It would also allow future studies to compare perceived collaboration potential with actual behaviours such as discussion quality, revision of ideas, agreement on a shared direction, or division of tasks.

### Limitations

Several limitations are important. First, this is a small pilot from one institution. There are 16 survey responses and only 14 valid collaboration responses, so individual answers have a large effect on the estimates. Second, collaboration potential is represented by one comparative item. A single item cannot establish a reliable multi-item construct, and the increase in Cronbach's alpha after removing it does not prove structural independence. Third, the study used a single post-use session and self-report data. Students did not complete a controlled collaborative task, so the paper cannot make claims about actual teamwork, learning gains, coordination, negotiation, or project quality. Fourth, several correlations were inspected in a small dataset. The p-values should therefore be read as descriptive rather than confirmatory. Fifth, the same

dataset has been used in earlier publications. We address this by keeping the present contribution narrow and by explicitly separating it from the earlier broad evaluation and design-prioritisation analyses. The new contribution is the the pattern of associations involving collaboration potential, not a new claim that the tool was generally well received.

## Conclusion

Perceived collaboration potential received positive ratings in this EduIDEAtor pilot, but it did not behave like most of the other quality ratings. Engagement, creativity, ease of use, and project relevance generally moved together. Collaboration potential had weak or near-zero associations with engagement, creativity, and ease of use, and a clearer association only with project relevance. Removing the collaboration item also increased the internal consistency of the remaining comparative block. With 14 valid responses and a single collaboration item, this should be treated as an exploratory signal rather than evidence for a validated separate construct. The practical value of the result is that it points to a question that broad usability scores can hide. A tool can feel easy and creative without telling us whether its outputs work well in a group. Future studies should therefore measure collaboration directly, using group tasks and collaboration-specific features, and compare perceived shareability with what students actually do together.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

OpenAI's ChatGPT was used during revision to improve language, readability, and organisation. All numerical analyses, interpretations, references, and final wording were checked and approved by the authors. Generative AI was not used to create or modify the study data.

## References

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist, 26*(3-4), 369–398. https://doi.org/10.1080/00461520.1991.9653139

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1–19). Elsevier.

Farokhi, K., & Zourlidou, S. (2026). EduIDEAtor survey: Perceptions of an LLM-based idea development tool [Data set]. Zenodo. https://doi.org/10.5281/zenodo.18171000

Farokhi, K., Zourlidou, S., & Hopfgartner, F. (2026). Does a minimal LLM assistant help students start? A pilot evaluation of an LLM-based ideation tool. In *INTED2026 Proceedings* (Article 1134). IATED. https://doi.org/10.21125/inted.2026.1134

Frich, J., MacDonald Vermeulen, L., Remy, C., Biskjaer, M. M., & Dalsgaard, P. (2019). Mapping the landscape of creativity support tools in HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Article 389, pp. 1–18). Association for Computing Machinery. https://doi.org/10.1145/3290605.3300619

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410

Shneiderman, B. (2007). Creativity support tools: Accelerating discovery and innovation. *Communications of the ACM, 50*(12), 20–32. https://doi.org/10.1145/1323688.1323689

Zourlidou, S., Farokhi, K., & Hopfgartner, F. (2026). Improving an LLM project ideation tool: evidence from a pilot study. In L. Gómez Chova, C. González Martínez, & J. Lees (series eds.), *EDULEARN26 Proceedings: 18th International Conference on Education and New Learning Technologies, 29 June-1 July, 2026, Palma, Spain* (p. 462). IATED Academy. https://doi.org/10.21125/edulearn.2026.0462

**Contact email:** zourlidou@uni-koblenz.de

# Design Conditions for Educator Adoption of LLM-Based Ethics Scaffolding in Higher Education

Stefania Zourlidou, University of Koblenz, Germany

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This paper examines what has to align before educators can integrate LLM-based ethics scaffolding into ordinary higher-education teaching without increasing marking burden, rule ambiguity, or inappropriate reliance on automated output. The paper develops an educator-facing framework from a bounded cross-case synthesis of recent higher-education literature and two prior design cases, one addressing classroom use of an LLM-supported ethics activity, and one addressing assessment redesign for keeping student reasoning inspectable under AI use. The paper proposes six design conditions for course integration: pedagogical positioning, workflow fit, visible reasoning, governance clarity, infrastructural reliability, and educator capability-building. These conditions suggest that adoption is less a matter of individual willingness to use AI than a matter of coordinated design across pedagogy, assessment, policy, infrastructure, and staff development. The paper contributes a practical framework, a temporal reading of how the conditions surface across the life of a course, and a worked illustration showing how the framework travels beyond computing to other judgment-intensive disciplines. Its central argument is that responsible adoption becomes more plausible when AI is bounded to a clearly defined pedagogical role and when student deliberation remains interpretable, assessable, and governable in everyday teaching conditions.

*Keywords:* generative AI, higher education, educator adoption, ethics education, assessment design

iafor
The International Academic Forum
www.iafor.org

## Introduction

Generative AI has not eliminated the educational problem of ethics teaching, instead, it has made course-design choices more complicated. In many higher-education settings, instructors are under pressure to respond to tools that can produce fluent analyses on demand, although the educational target in ethics-related work is rarely fluent text alone. It is the capacity to frame a dilemma, weigh competing considerations, justify a position under uncertainty, and revise that position when better evidence or criticism appears. When AI tools enter this space, the practical question for educators is not simply whether a model can be useful, but whether the surrounding activity still keeps judgment visible, responsibilities clear, and marking manageable. That question extends well beyond computing courses, arising anywhere students are asked to reason through contested cases, professional responsibilities, or uncertain evidence. This paper addresses that problem by developing an educator-facing framework for course integration. It draws on a bounded cross-case synthesis of recent post-2022 literature on AI in higher education together with two prior design cases from the same broader research programme. The first is an exploratory classroom pilot of EthicsAdvisor (Oloke & Zourlidou, 2026), a structured LLM-based scaffold for ethical reflection in data-science projects. The second is an assessment-design study that introduced an Ethics Decision Log (EDL), a compact reasoning rubric, and a critique-and-revision workflow for AI-rich project courses (Zourlidou et al., 2026). The present paper does not re-evaluate EthicsAdvisor, reanalyse the pilot participants, or present the EDL and rubric as new contributions. Instead, the two prior cases serve as bounded analytic inputs to a different question: what has to align before an educator can adopt this kind of scaffold as a credible part of ordinary teaching practice? The resulting framework is aimed at educators and programme teams who need a practical way to think about adoption across pedagogy, assessment, governance, infrastructure, and professional learning.

## Related Work

Recent reviews make clear that AI in higher education is already a large and fast-moving field, while also showing why an educator-facing synthesis remains useful. Bond et al. (2024), in a meta-review of secondary research on AI in higher education, identify rapid growth alongside uneven ethical, methodological, and contextual grounding. Their synthesis also shows substantial attention to broad AI-in-higher-education questions and to profiling and prediction. Crompton and Burke (2023), reviewing primary research from 2016 to 2022, identify assessment and evaluation, prediction, AI assistants, intelligent tutoring systems, and the management of student learning among prominent applications, while students receive more research attention than instructors. Read together, these reviews show a field rich in accounts of AI applications but less developed in educator-centred accounts of how such tools can be integrated responsibly into ordinary teaching practice. The present paper responds to that gap by shifting the level of analysis from technical possibility and student reception to adoption conditions for instructors.

A second body of work shows why educator adoption cannot be reduced to individual openness or technical skill. Chan (2023) proposes an AI Ecological Education Policy Framework organized across pedagogical, governance, and operational dimensions, explicitly linking teaching and learning to privacy, security, accountability, infrastructure, and training. Dotan et al. (2024) argue that higher education's traditions of collaborative faculty governance and academic freedom can sit uneasily with centralized, top-down approaches to AI governance. Jin et al. (2025), analysing policies and guidelines from 40 universities across six global regions, find proactive attention to academic integrity, teaching and learning, and equity, while

identifying remaining gaps particularly in data privacy and equitable access. They also emphasize clear communication channels, stakeholder collaboration, and defined roles and responsibilities. Miao and Holmes (2023) similarly place GenAI policy within a human-centred approach that foregrounds data privacy, human agency, and ethical and pedagogical validation. Therefore these findings suggest that responsible integration depends not only on educator willingness but also on policy clarity, institutional support, and workable governance arrangements.

The assessment literature sharpens the problem by showing that adoption is inseparable from what educators are being asked to grade. Luo (2024), analysing GenAI policies at 20 universities, found that the dominant policy framing centred on the originality of students' submitted work, even though other issues such as reliability and equal access were also present. Smolansky et al. (2023) and Kizilcec et al. (2024) report that educators strongly favour assessment approaches adapted to assume AI use and to encourage critical thinking, whereas student responses are more mixed. Khlaif et al. (2025) identify faculty motivations for assessment redesign that include academic integrity, preparation for future careers, technological change, and institutional policy, while also reporting professional-development, equity, and accessibility challenges. Xia et al. (2024) conclude that assessment transformation should cultivate self-regulated learning, responsible learning, and integrity, and should be accompanied by teacher professional development and more innovative, holistic teaching. This literature supports a design requirement central to the present synthesis, i.e., if an AI-related scaffold is to become part of routine teaching, the assessed work must leave interpretable evidence of student thinking and not rely on polished output alone.

A final strand of research brings educator capability and professional learning into view. Mah and Groß (2024) identify distinct faculty profiles associated with differences in AI use, self-efficacy, perceived benefits and challenges, and professional-development needs, with AI literacy emerging as an important challenge. Shahzad et al. (2025), in qualitative interviews with 29 management educators, report concerns including ethics, data privacy, academic integrity, institutional readiness, and the need to balance automation with human-centred teaching. Miao and Cukurova (2024) organize teacher AI competence across five dimensions: human-centred mindset, ethics of AI, AI foundations and applications, AI pedagogy, and AI for professional learning. This literature helps distinguish between visible reasoning, which concerns what an assignment makes observable in student work, and capability-building, which concerns what educators need to understand and be able to do to interpret that work, set boundaries for tool use, and teach students to evaluate model outputs critically.

## Methodology

Methodologically, this paper is an integrative design synthesis, not a systematic review or a new empirical instructor-readiness study. The goal is to build a practically usable framework for one specific problem: what conditions make educator adoption of LLM-based ethics scaffolding plausible in ordinary course settings. This orientation follows integrative-review methodology, which uses existing literature to create new knowledge through synthesis (Torraco, 2016). The external literature therefore functions here as framework-building evidence. The corpus was limited to recent sources from 2023 onward that speak directly to higher-education AI adoption, governance, assessment redesign, educator perspectives, or professional capability. The two prior cases were deliberately read at the level of function and not treated as new empirical data. The EthicsAdvisor case was an exploratory, single-site classroom pilot with 20 data-science students that reported perceptions of usefulness, clarity,

applicability, and related classroom experience (Oloke & Zourlidou, 2026). In the present synthesis, that published case is used only to surface implementation questions such as pedagogical role framing, low-friction access, response behaviour, and the practical consequences of service latency or verbosity. No participant-level data were reanalysed, and the pilot's student-perception results are not presented as findings about educator adoption. The second case developed an assessment package consisting of the EDL, a 0-2 reasoning rubric, critique prompts, and worked exemplars (Zourlidou et al., 2026). Here, that case is used to examine the assessment implications of keeping student judgment inspectable under AI use. The present paper does not reproduce the EDL template or rubric as a new artifact, and it makes no new claim about their effectiveness.

The synthesis proceeded in four passes. First, the two prior cases were compared for the educator-facing problems they exposed at different points in a course. Second, each external source was coded for the educator-facing issue it illuminated, including workload, policy ambiguity, integrity, assessment redesign, staff capability, and institutional support. Third, recurrent concerns were clustered across the literature and tested against the two cases to determine whether they appeared across different parts of the teaching process rather than as isolated themes. Fourth, the clusters were refined into six design conditions, each expressed as an educator-facing question and linked to a course-integration implication (Table 1).

This procedure has clear limitations. The synthesis is deliberately bounded, the two prior cases are close to one another in topic and institutional context, and one is exploratory (not large-scale). The paper therefore does not claim that the six conditions are exhaustive, that all instructors experience them in the same way, or that the framework has already been validated across disciplines. Its more modest claim is that when recent literature is read alongside bounded course-design cases, a coherent set of educator-facing alignment problems becomes visible. The framework organizes those problems for further testing and for practical course-design discussion.

**Table 1**
*Six Design Conditions for Educator Adoption*

| Condition | Educator-facing question | Course integration implication |
|---|---|---|
| Pedagogical positioning | What precise teaching job is the scaffold being asked to do? | Define the scaffold as a bounded aid for deliberation, not a substitute for judgment. |
| Workflow fit | Can the activity fit ordinary teaching time and grading time? | Use low-friction access, short task cycles, and explicit contingency steps. |
| Visible reasoning | Can student judgment be seen without intrusive monitoring? | Assess claims, warrants, revisions, and evidence checks rather than polished prose alone. |
| Governance clarity | Do rules for use, disclosure, privacy, and retention make sense to everyone? | Provide concise, transparent expectations tied to pedagogical purpose. |
| Infrastructural reliability | Will the tool behave predictably enough for live coursework? | Prepare for service instability, response-length control, and non-AI continuation routes. |
| Educator capability-building | Does the instructor know how to teach with the scaffold and interpret student work? | Support staff with calibration opportunities and professional learning on ethics, equity, and critique. |

## Framework and Synthesis Findings

The synthesis identifies pedagogical positioning as the first condition. Before a tool is evaluated for convenience or novelty, the educator has to decide what educational role it is allowed to play. Across the reviewed literature and the two prior cases, integration appears more defensible when the scaffold is bounded to a narrow judgment-supporting function: helping students surface relevant considerations, test assumptions, or widen a line of inquiry without claiming authority over the conclusion itself. This positioning is consistent with the EthicsAdvisor case, where model output was framed as scaffolding rather than as a decision or authoritative answer (Oloke & Zourlidou, 2026). The precise domain may differ, fairness in computing, confidentiality in nursing, professional responsibility in law, or stakeholder accountability in business, but the pedagogical issue is constant: the tool should support deliberation without displacing responsibility.

The second condition is workflow fit. A practice may be conceptually attractive and still be unsuitable for routine course use if it cannot live inside ordinary teaching and grading time. The synthesis repeatedly points to low-friction access, manageable task length, predictable review demands, and explicit contingency steps as relevant integration concerns. The earlier classroom pilot also documented first-call latency and occasional verbosity as practical limitations, illustrating how apparently minor interaction costs can shape the feasibility of live use (Oloke & Zourlidou, 2026). Workflow fit therefore concerns both classroom activity and marking time: a sound pedagogical idea can still fail as a course practice if it is cumbersome to run, supervise, or assess at scale.

The third condition is visible reasoning, the point at which pedagogical aims and assessment requirements meet. For adoption to be credible, the activity needs to leave an interpretable trail of how students framed a case, weighed competing considerations, checked claims, and revised their position. The prior assessment-design study operationalized this concern through an EDL and rubric that grade decision-specific reasoning and revision rather than polished prose (Zourlidou et al., 2026). The present framework generalizes the design requirement rather than reintroducing those artifacts: visible reasoning can be implemented in different ways, but the work submitted for assessment should make claims, warrants, revisions, evidence checks, and uncertainty inspectable. When such a trail is absent, fluent AI-assisted prose can obscure shallow engagement; when it is present, educators have a basis for judging deliberative quality without relying on intrusive monitoring.

The fourth condition is governance clarity. Educators can often work around imperfect tools more easily than they can work around ambiguous rules. The reviewed literature points to governance questions concerning acceptable use, disclosure, privacy, equitable access, communication, and responsibility for AI-supported teaching and assessment (Dotan et al., 2024; Jin et al., 2025; Miao & Holmes, 2023; Shahzad et al., 2025). The two prior cases reinforce the practical importance of making expectations explicit. For course integration, clarity need not mean an elaborate policy; it can mean a concise explanation of why the scaffold is being used, what students must disclose, what data are retained, and what continuation route exists when AI is unavailable or inappropriate. Without such clarity, otherwise well-designed activities are harder to justify and harder for students to interpret consistently.

The fifth condition is infrastructural reliability. Although this condition may look technical, its consequences are pedagogical. If access is unstable, response length is unpredictable, or live sessions stall, the activity can move away from deliberate reasoning and toward frustration or

hasty acceptance. The classroom-use case provides a bounded illustration of this problem through observed latency and verbosity issues, while also showing that explicit limitations and verification advice can help keep model output in a supporting role (Oloke & Zourlidou, 2026). The synthesis therefore treats reliability broadly: service stability, predictable interaction cost, scope control, and a clear non-AI continuation route all matter. A tool does not need to be perfect to be teachable, but it does need to be dependable enough that an educator can plan around it.

The sixth condition is educator capability-building. This condition is distinct from visible reasoning even though the two interact closely. Visible reasoning concerns what the activity makes available for inspection; capability-building concerns what the instructor must know and be able to do in order to design the task, explain boundaries, interpret student work, and challenge inappropriate reliance on model output. Recent studies identify self-efficacy, AI literacy, ethical and privacy concerns, institutional readiness, and professional-development needs as salient factors in educators' experiences of AI adoption (Mah & Groß, 2024; Shahzad et al., 2025). The teacher competency framework of Miao and Cukurova (2024) further situates capability across human-centred, ethical, technical, pedagogical, and professional-learning dimensions. In practical terms, capability-building means that instructors need examples, calibration opportunities, and a language for explaining both the value and the limits of AI-supported critique. Without that support, even a well-designed activity can fail because the educator has been asked to improvise too much.

A brief course illustration shows how the six conditions can operate together. Consider a third-year course in which teams must analyse the ethical implications of a proposed AI-enabled service. The same underlying structure could be adapted for a computing lab, a nursing case seminar, a law-and-technology elective, or a business analytics capstone. The instructor introduces the AI scaffold as a bounded aid for testing assumptions and broadening analysis, not as a source of final answers. Students submit a short reasoning record at an intermediate checkpoint and a revised version later, explaining how their position changed, what evidence or critique informed that change, and where uncertainty remains. This checkpoint-and-revision pattern is informed by the earlier assessment-design case (Zourlidou et al., 2026), but it is used here only as an illustration of how the six adoption conditions can be coordinated, not as a newly evaluated intervention. The assessment focuses on the quality of the deliberative trail rather than stylistic polish; students receive clear guidance about acceptable use and disclosure; a non-AI fallback is available; and the instructor has seen calibrated examples before grading begins.

There is also a temporal logic worth making explicit. The six conditions do not merely coexist. Instructors encounter them at different points across the life of a course. Before teaching begins, pedagogical positioning and governance clarity dominate, because the educator has to decide what the scaffold is for, whether it is permitted, and how expectations will be communicated. When the task is introduced, workflow fit and capability-building become more visible, because students need manageable instructions and the instructor needs enough fluency to explain examples, boundaries, and likely failure modes. During active use, infrastructural reliability and visible reasoning come to the foreground, because access problems or over-fluent output can immediately distort participation and judgment. At grading time, visible reasoning, workflow fit, and capability-building converge again, since the educator must interpret evidence efficiently and distinguish substantive revision from cosmetic editing. This time-based reading strengthens the framework by treating adoption not as a one-time decision but as a chain of design commitments that must survive contact with teaching practice.

## Discussion

The main implication of the synthesis is that educator adoption is best understood as a coordination problem and not as a question of individual attitude or technical enthusiasm. The framework is not intended to prescribe a single artifact, interface, or workflow but its purpose is to help educators judge whether local course conditions are sufficient for responsible integration. An instructor may recognize pedagogical value in AI-supported ethics work and still avoid adoption if the workflow is too cumbersome, if institutional rules are vague, or if the platform cannot be relied on during class. Conversely, a technically smooth tool may still fail educationally if it obscures the deliberative work that assessment needs to see or shifts authority toward model fluency. This is why the six conditions should not be read as an additive checklist in which strength in one area easily compensates for weakness in another. Pedagogical positioning gives the activity a legitimate teaching purpose. Workflow fit makes it livable, visible reasoning makes it reviewable, governance clarity makes it justifiable, infrastructural reliability makes it dependable in teaching time, and capability-building makes it teachable.

For institutions and programme teams, the framework suggests a distributed model of support. Adoption should not be delegated entirely to the individual educator, but neither can it be solved by generic top-down policy statements alone. Programme leaders can identify where judgment-heavy tasks already appear in capstones, case seminars, simulations, or reflective placements, learning designers can help reshape those moments into assessable activities, IT teams can support low-friction access and continuation routes, and academic development units can provide exemplars, calibration routines, and professional learning that address ethics and equity as well as tool operation. This institutional view matters beyond computing. A business lecturer redesigning a consultancy brief, a nursing educator structuring case-based reflection, or a law instructor using AI-supported critique in a professional-responsibility seminar may not share the same disciplinary standards, but they face similar adoption questions about clarity, trust, workload, and student accountability. The framework is therefore best read as discipline-sensitive but structurally portable.

The limitations of the paper should remain visible. The framework is grounded in a bounded corpus and two prior design cases that are close to one another in topic and institutional context, and one of those cases is exploratory rather than large-scale. The synthesis is therefore better at identifying plausible design conditions than at ranking them by causal importance. It also does not settle discipline-specific questions about what counts as sufficient evidence, feasible mitigation, acceptable disclosure, or sustainable workload in every field. Those judgments still require local interpretation. Future work should test the six conditions across a broader disciplinary range, examine how they interact with institutional policy maturity, and explore how staff capability-building and visible reasoning reinforce one another over time.

## Conclusion

This paper proposes six aligned design conditions for educator adoption of LLM-based ethics scaffolding in higher education: pedagogical positioning, workflow fit, visible reasoning, governance clarity, infrastructural reliability, and educator capability-building. The article develops an educator-facing framework from a bounded cross-case synthesis of recent literature and two prior design cases. In doing so, it shifts attention from whether a model or artifact works in isolation to the educational and institutional alignments required for credible course integration. The broader lesson is that responsible adoption should make human judgment more interpretable, not less. In AI-rich classrooms, educators need activities,

interfaces, and support structures that keep deliberation open to review, revision meaningful, and accountability human. When an AI scaffold is given a bounded pedagogical role, embedded in a manageable workflow, supported by clear rules, delivered through dependable infrastructure, accompanied by assessable evidence of student deliberation, and matched by adequate educator learning, it becomes more plausible as a support for ethical reflection without displacing responsibility. Although the argument begins with ethics-related teaching, its relevance extends to other higher-education settings in which students are asked to reason through contested cases, professional obligations, and uncertain evidence.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that OpenAI ChatGPT (GPT-5.6 Sol) was used to assist with language refinement, structural editing, and manuscript formatting. The tool was not used to generate, alter, or analyse study data. The author reviewed the resulting text and retains responsibility for the argument, interpretation, citations, and final submitted version.

## References

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education, 21*, Article 4. https://doi.org/10.1186/s41239-023-00436-z

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education, 20*, Article 38. https://doi.org/10.1186/s41239-023-00408-3

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*, Article 22. https://doi.org/10.1186/s41239-023-00392-8

Dotan, R., Parker, L. S., & Radzilowicz, J. (2024). Responsible adoption of generative AI in higher education: Developing a "points-to-consider" approach based on faculty perspectives. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 2033–2046). Association for Computing Machinery. https://doi.org/10.1145/3630106.3659023

Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence, 8*, 100348. https://doi.org/10.1016/j.caeai.2024.100348

Khlaif, Z. N., Alkouk, W. A., Salama, N., & Abu Eideh, B. (2025). Redesigning assessments for AI-enhanced learning: A framework for educators in the generative AI era. *Education Sciences, 15*(2), 174. https://doi.org/10.3390/educsci15020174

Kizilcec, R. F., Huber, E., Papanastasiou, E. C., Cram, A., Makridis, C. A., Smolansky, A., Zeivots, S., & Raduescu, C. (2024). Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States. *Computers and Education: Artificial Intelligence, 7*, 100269. https://doi.org/10.1016/j.caeai.2024.100269

Luo, J. (2024). A critical review of GenAI policies in higher education assessment: A call to reconsider the "originality" of students' work. *Assessment & Evaluation in Higher Education, 49*(5), 651–664. https://doi.org/10.1080/02602938.2024.2309963

Mah, D. K., & Groß, N. (2024). Artificial intelligence in higher education: Exploring faculty use, self-efficacy, distinct profiles, and professional development needs. *International Journal of Educational Technology in Higher Education, 21*, Article 58. https://doi.org/10.1186/s41239-024-00490-1

Miao, F., & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO. https://doi.org/10.54675/ZJTE2084

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Oloke, P., & Zourlidou, S. (2026). EthicsAdvisor: A pilot LLM-based tool for ethical reflection in data science education. In *2026 22nd International Conference on Information Technology Based Higher Education and Training (ITHET),* Lilehammer, Norway, 2026, pp. 1–8, IEEE. https://doi.org/10.1109/ITHET69978.2026.11584968

Shahzad, F., Avunduk, Z. B., Ullah, Z., & Arslan, A. (2025). Understanding management educators' perspectives on AI adoption: A qualitative study. *The International Journal of Management Education, 23*(3), 101274. https://doi.org/10.1016/j.ijme.2025.101274

Smolansky, A., Cram, A., Raduescu, C., Zeivots, S., Huber, E., & Kizilcec, R. F. (2023). Educator and student perspectives on the impact of generative AI on assessments in higher education. In *Proceedings of the Tenth ACM Conference on Learning @ Scale* (pp. 378–382). Association for Computing Machinery. https://doi.org/ 10.1145/3573051.3596191

Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. *Human Resource Development Review, 15*(4), 404–428. https://doi.org/10.1177/1534484316671606

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education, 21*, Article 40. https://doi.org/10.1186/s41239-024-00468-z

Zourlidou, S., Oloke, P., & Hopfgartner, F. (2026). Grading the reasoning, not the answer: Assessable artifacts for LLM-assisted ethics in data science projects. In *Proceedings of the 18th International Conference on Computer Supported Education (CSEDU 2026)* (Vol. 2, pp. 1578–1589). SciTePress. https://doi.org/10.5220/0014826600004021

**Contact email:** zourlidou@uni-koblenz.de

# 3 Level Game for the First-Year Students in Computer Science

Daniela Tsaneva, Cardiff University, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

This study explores the use of gamification and Generative AI to enhance engagement and learning outcomes to "*Design, build, and test front-end web pages according to specifications and best practices"* in introductory web development education. It addresses the research question: "*To what extent can a gamified, AI-supported learning intervention improve first-year students' understanding of HTML and CSS while promoting meaningful engagement?"* A three-level HTML & CSS Mastery Game was designed and implemented within a Web Applications module at Cardiff University. It comprises: (1) *Builder Basics*, focusing on HTML structure through interactive drag-and-drop tasks; (2) *CSS Styling*, requiring students to complete code based on design prompts; and (3) *Guess the Tag*, reinforcing conceptual understanding through matching activities. Generative AI contributed to the pedagogical design by supporting the creation of varied task instances, generating immediate and adaptive feedback, and enabling scalable content development. Its use was guided by principles of responsible AI, ensuring transparency, accuracy, and alignment with educational objectives. A mixed-methods research approach was adopted with 260 first-year students, combining quantitative performance data with qualitative feedback from 73 questionnaire responses. Findings indicate increased student engagement, improved conceptual understanding, with participants reporting higher motivation and a preference for the gamified approach compared to traditional formative assessments. This study contributes to the field of higher education pedagogy by presenting a scalable framework for integrating gamification and Generative AI into computing education. It highlights how AI can be used to enhance active learning and formative assessment while maintaining pedagogical integrity and supporting inclusive education.

*Keywords:* gamification, generative AI, formative assessment

iafor
The International Academic Forum
www.iafor.org

## Introduction

Gamification in higher education has an important role in boosting student motivation, engagement, and cognitive performance. Applying gamification in formative assessment leads to increased student engagement, improved conceptual understanding, higher motivation and a preference for the gamified approach compared to traditional formative assessments.

In recent days Artificial Intelligence (AI) has been employed in every area of science, business and everyday life. Therefore, It is essential to employ the assistance of Generative AI in designing learning materials and assessments to support the creation of varied task instances, generate immediate and adaptive feedback, and enable scalable content development.

## Literature Review

According to Ortiz et al. (2016) in recent years, gamification, the use of game elements in non-game contexts, has drawn the attention of educators due to the possibility of making learning more motivating and engaging. Analysis results show how a combination of game elements (e.g. leaderboards, badges, points and other combinations) positively affects students' performance, attendance, goal orientation and attitude towards mostly computer science related subjects.

The study in Gyedu et al. (2026) highlights the importance of context-specific and theory-driven gamification strategies tailored to learners' needs and institutional environments. In conclusion, findings affirm that when carefully implemented, gamification is a powerful tool for improving learning outcomes in higher education.

The results from the research conducted in Camacho et al. (2026) show that the integration of game mechanics —such as points, levels, challenges, and rewards— increases student motivation and commitment, facilitates the development of critical thinking, communication, and teamwork, and redefines the role of the teacher as a designer of training experiences.

Therefore, there is a need to explore the use of gamification and Generative AI to enhance engagement and learning outcomes in introductory web development educatiojn. This study contributes to the field of higher education pedagogy by presenting a scalable framework for integrating gamification and Generative AI into computing education.

Nowadays, Artificial Intelligence (AI) has a dramatic impact on our everyday lives and in particular on Higher Education (HE). Therefore, since universities need to prepare the students for the industry, it is extremely important that the HE sector stays informed of developments in AI and commit to ongoing research and discussions regarding best practice in accommodating AI in learning and assessment.

Besides mentioning challenges for HEIs, some authors also see opportunities for student learning, feedback, and assessment design. Bobula (2024) emphasises the positive role AI can play in education.

Furthermore, AI-supported educational tools enable personalized learning by identifying individual students' strengths and weaknesses and providing precise performance assessments as mentioned by Chen et al. (2022). The authors in Francis et al. (2025) propose a set of

recommendations for educational institutions, including developing GenAI literacy programmes and revising assessment designs to incorporate critical thinking and creativity.

### Generative AI – ChatGPT

ChatGPT is an AI language model developed by OpenAI that is designed to understand and generate human-like text based on the input it receives. It uses machine learning, particularly a type of neural network called a Transformer, to generate coherent and contextually relevant responses. Trained on vast amounts of text data from various sources, it can assist with a wide range of tasks, including answering questions, providing explanations, generating creative content, coding, tutoring, and more.

Key features of ChatGPT include:

- ***Natural Language Processing (NLP):*** ChatGPT can understand and generate text in a conversational style, making interactions feel more human-like.
- ***Wide Range of Topics***: It is knowledgeable in many areas, such as science, history, technology, arts, entertainment, and much more.
- ***Context Awareness***: ChatGPT tries to maintain context during a conversation, allowing for follow-up questions and more dynamic interactions.
- ***Customization and Use:*** It can be embedded in various applications or used for specific tasks like writing, customer service, and content creation.
- ***Limitations:*** While powerful, it may produce incorrect information, lacks real-time awareness, and has no understanding or awareness of personal context unless specifically shared during the conversation.

It is like an advanced chatbot with diverse applications in customer support, education, entertainment, and productivity.

## Research Question

This study explores the use of gamification and Generative AI to enhance engagement and learning outcomes to "Design, build, and test front-end web pages according to specifications and best practices" in introductory web development education. It addresses the research question: To what extent can a gamified, AI-supported learning intervention improve first-year students' understanding of HTML and CSS while promoting meaningful engagement?

The First-year students in Computer Science at Cardiff University have been enjoying active learning based on Generative AI methods of teaching and assessment. This study contributes to the field of higher education pedagogy by presenting a scalable framework for integrating Gamification and Generative AI into computing education. It highlights how AI can be used to enhance active learning and formative assessment while maintaining pedagogical integrity and supporting inclusive student engagement.

## Methodology

### Formative Assessment Through Gamification

Learning how to do web programming and design web sites can be sometimes too dry and boring, The First-Year students in Computer Science at Cardiff University found a more interesting way of learning through Gamification. Their knowledge about the Web

Applications module would be usually evaluated during the in-person sessions through Mentimeter and quizzes through which we can get some feedback about their learning, but it is not very specific and always follows the same format.

Therefore, we have designed with the help of Generative AI a simple HTML&CSS Mastery Game with 3 levels to evaluate the students' skills in web design.

Generative AI contributed to the pedagogical design by supporting the creation of varied task instances aligned with learning outcomes, generating immediate and adaptive feedback, and enabling scalable content development. Its use was guided by principles of responsible AI, ensuring transparency, accuracy, and alignment with educational objectives.

## Game Implementation

### Three-Level HTML & CSS Mastery Game

A three-level HTML & CSS Mastery Game was designed and implemented within a Web Applications module at Cardiff University for the 1st year students in Computer Science. The game comprises of:

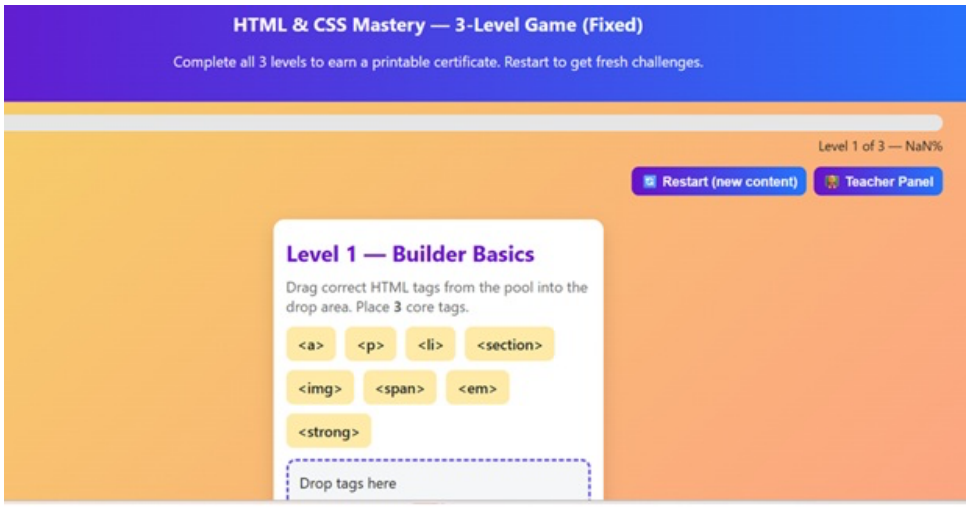
1. ***Builder Basics*****,** focusing on HTML structure through interactive drag-and-drop tasks;
2. ***CSS Styling*****,** requiring students to complete code based on design prompts; and
3. ***Guess the Tag*****,** reinforcing conceptual understanding through matching activities.

Feedback is provided for correct and incorrect answers. In addition, there is a progress bar during each level showing how much the student has achieved. Finally, a certificate is rewarded for the successful completion of the game.

### Level 1: Builder Basics

Level 1 is called Builder Basics and asks the students to drag the correct HTML tags from the pool provided into the drop area. The students are supposed to place 3 core tags and are provided with feedback which tags are correctly or incorrectly placed. An example of the Level 1 is presented in **Figure 1** below.

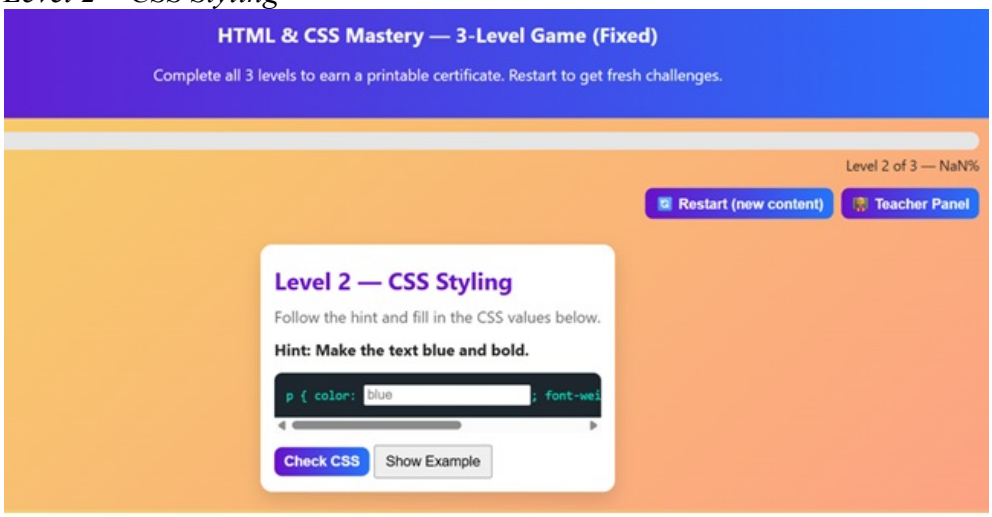
**Figure 1**
*Level 1 – Builder Basics*




## Level 2 – CSS Styling

In Level 2 CSS Styling the students need to follow the hint and fill in the CSS values in the code below. (e.g. Hint: Make the text blue and bold). The students can check the correct CSS and are provided with feedback. An example of the Level 2 is presented in **Figure 2** below.

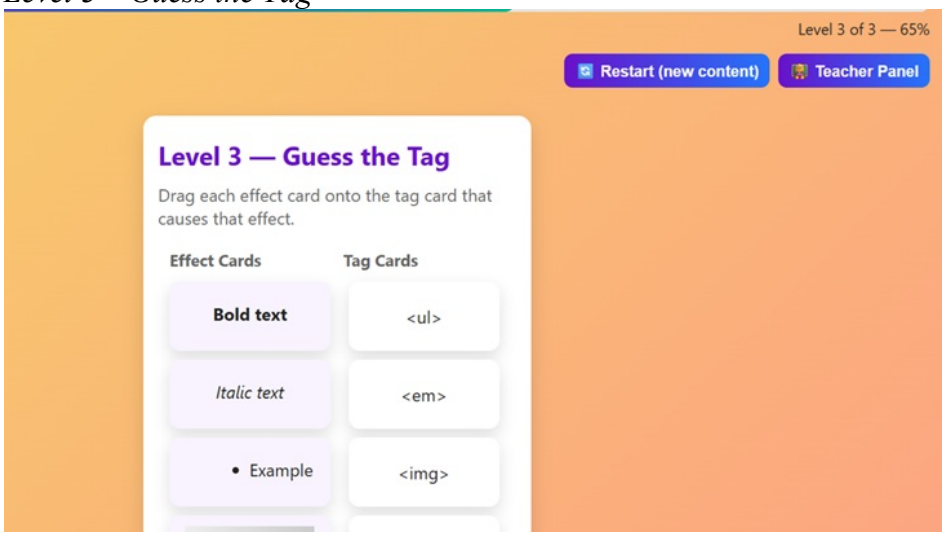
**Figure 2**
*Level 2 – CSS Styling*




## Level 3 – Guess the Tag

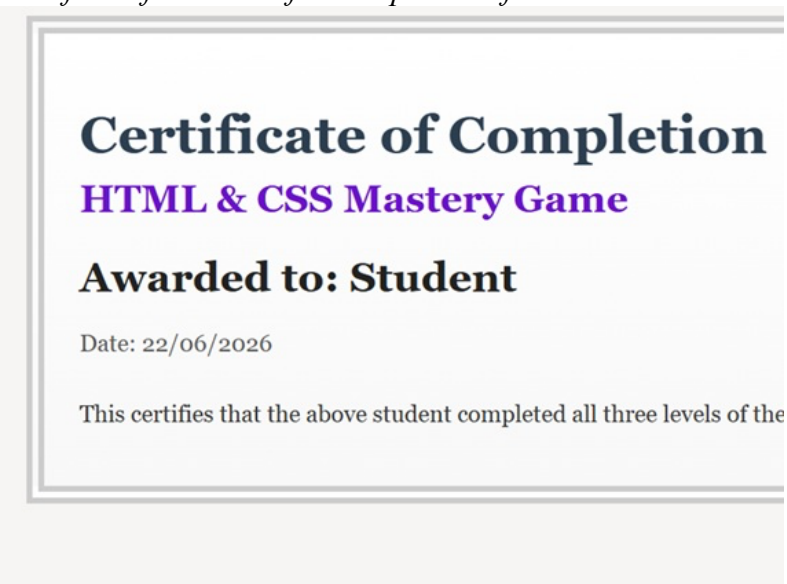
In Level 3 Guess the tag the students need to drag each effect card onto the tag card that causes that effect. The students can check the correct tag and are provided with feedback. An example of the Level 3 is presented in **Figure 3** below.

**Figure 3**
*Level 3 – Guess the Tag*




On the successful completion of the game the students receive a certificate with their name for the completion of the HTML and CSS mastery game. An example of the certificate is presented in **Figure 4** below.

**Figure 4**
*Certificate for Successful Completion of the Game*




## Results and Discussion

A mixed-methods research approach was adopted with 260 first-year students, combining quantitative performance data with qualitative feedback from 73 questionnaire responses. The students' response was collected through their feedback to a Mentimeter quiz designed for that purpose. Findings indicate increased student engagement and understanding, with participants reporting higher motivation and a preference for the gamified approach compared to traditional formative assessments.

Key insights include:

- The game felt easy, but many still said they learned little, so the challenge may have been too low for retention.
- Hints and answer feedback were only partly clear, suggesting the core tags step needs simpler wording and examples.
- Enjoyment was generally positive, but confusion around “main” and “core” tags show the first level needs a clearer setup.

## Conclusions and Future Work

This study contributes to the field of higher education pedagogy by presenting a scalable framework for integrating gamification and Generative AI into computing education. It highlights how AI can be used to enhance active learning and formative assessment providing pedagogical integrity and supporting inclusive education.

The key insights from the students’ feedback show that the game needs improvement. After conducting some further research into gamification additional ideas for future formative assessment have been generated, including online Escape room and a Quest game or Interactive game based on real-life scenario and AI song creation with lyrics based on Computer Science.

## References

Bobula, M. (2024). Generative Artificial Intelligence (AI) in higher education: a comprehensive review of challenges, opportunities, and implications. *Journal of Learning Development in Higher Education,* (30), 112–127. https://doi.org/10.47408/jldhe.vi30.1137

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology & Society, 25*(1), 28–47. https://www.jstor.org/stable/48647028

Francis, N. J., Jones, S., & Smith, D. P. (2025). Generative AI in higher education: Balancing innovation and integrity. *British Journal of Biomedical Science, 81,* Article 14048. https://doi.org/10.3389/bjbs.2024.14048

Gyedu, F. O., Partey, P. A., Boafo, F. A., & Agbo, D. D. (2026). Gamification effect: Enhancing student learning outcomes in higher education: A meta-analysis and policy implications. *SAGE Open, 16*(1), 1–21. https://doi.org/10.1177/21582440261421375

Ortiz Rojas, M. E., Chiluiza, K., & Valcke, M. (2016). Gamification in higher education and STEM: A systematic review of literature. In *EDULEARN16 proceedings: 8th International Conference on Education and New Learning Technologies* (pp. 6548–6558). IATED. https://doi.org/10.21125/edulearn.2016.0422

Quenema Camacho, N., Otero Bocanegra, P. C., Condeso Camizan, S. D., Hernández Ramos, E. J., & Castillo Palacios, F. W. (2026). Gamificación en la evaluación educativa: transformando el aprendizaje en experiencias interactivas desde una revisión sistemática [Gamification in Educational Assessment: Transforming Learning into Interactive Experiences through a Systematic Review]. *Revista InveCom, 6*(1), 1–10. https://zenodo.org/records/15427281

# Evaluating Simulation-Based Learning in Management Education: Evidence From Student Perceptions

Salman Ali, Indian Institute of Management Kozhikode, India

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Simulation-based pedagogy is increasingly used in management education to create experiential learning environments that connect theoretical concepts with real-world decision-making. This study examines the effectiveness of classroom simulations as a teaching tool in a postgraduate management course, with a focus on student engagement, learning outcomes, and skill development. The study draws on a classroom implementation involving postgraduate students. Data were collected using a structured post-simulation survey comprising Likert-scale items and qualitative reflections. The survey captured perceptions of conceptual understanding, application of theory, engagement, collaboration, and higher-order skills, while qualitative responses provided contextual insights into the learning process. We received 50 responses. The findings demonstrate consistently strong positive outcomes. Over 80% of students reported improved conceptual understanding and application of theory, and more than 85% indicated enhanced strategic decision-making and critical thinking. Engagement emerged as a key strength, with approximately 88% of students reporting higher engagement than in traditional lecture-based formats and active participation in simulation activities. Around 80% also highlighted improved collaboration and teamwork. These results are based on self-reported perceptions rather than objective assessments of learning, which limits claims regarding conceptual understanding and problem-solving outcomes. Nevertheless, qualitative evidence corroborates the findings, emphasizing the integration of functional knowledge and the translation of strategy into action. The study contributes to the literature by demonstrating the pedagogical value of simulations in fostering experiential, applied, and collaborative learning in management education.

*Keywords:* simulation-based learning, experiential learning, management education, business simulation, MBA

iafor
The International Academic Forum
www.iafor.org

## Introduction

Simulation-based pedagogy immerses students in realistic, decision-oriented environments, aligning with Kolb's (1984) experiential learning cycle (Kuczera, 2021). Kolb emphasizes learning through concrete experience and active experimentation, mirroring the hands-on nature of management simulations. Management educators argue that such simulations allow students to "experience" business dynamics and "learn by managing a simulated firm," thereby bridging theory and practice (Clarke, 2010). Prior literature highlights multiple benefits: simulations increase engagement and motivation, replicate complex market conditions for problem-solving, and allow learners to test strategies with immediate feedback (Keys & Wolfe, 1997). For example, Clarke (2010) found that simulations replicate dynamic economic events, helping students understand interrelationships among business functions and develop intuition and analytical skills in ill-structured contexts. It was observed that simulation environments enable instructors to directly "observe learning and behavioural changes," as students experiment with decisions in real time.

Extensive research confirms these claims. Salas et al. (2009) review evidence that simulation-based training (SBT) offers "many advantages" in management education and provide guidelines for its use. Velez and Alonso (2025) report that business simulation games (BSGs) are "considered a useful instructional tool" for fostering higher-order skills like decision-making. Their systematic review found that the majority of simulation experiences enhanced decision-making results for students. However, most published studies rely on self-reported surveys, and there is a need for more empirical data on actual classroom implementations. This study responds by documenting student perceptions following a specific course-based simulation.

### Objectives and Research Questions

The goal of this study is to evaluate student learning experiences and outcomes from a simulation activity in a postgraduate management course. Guided by experiential learning theory, we examine: (1) Do students perceive improvements in their conceptual understanding and ability to apply management theory? (2) How does the simulation affect engagement and team collaboration compared to lectures? (3) What qualitative insights do students report about their learning, including challenges and the overall value of the simulation? By combining quantitative and qualitative survey data, we aim to identify the strengths and limitations of our simulation-based pedagogy and provide practical implications for management curriculum design.

## Methodology

### Participants and Context

50 students from a class of 120 students (two sections of 60 each) in a postgraduate management strategy simulation course participated in this study. The cohort was mixed-gender and included students from diverse professional backgrounds. All participants engaged in a multi-week business strategy simulation (a competitive, web-based simulation managing a virtual firm across marketing, finance, and operations functions). The simulation required students to make periodic strategic decisions and allowed observation of outcomes over time.

## Procedure

After completing the simulation, students were invited to complete an anonymous online survey. The survey contained 10 Likert-scale items (1 = Strongly Disagree to 5 = Strongly Agree) and 3 open-ended questions. Likert items were adapted from established experiential learning and simulation outcome measures, targeting key constructs: conceptual understanding (e.g., "The simulation helped me understand course concepts more clearly"), application of theory ("The simulation helped me apply theoretical knowledge"), decision-making ("The simulation improved my decision-making skills"), critical thinking, engagement, and team collaboration. The open-ended prompts asked students to describe the most valuable aspects of the simulation, any challenges they faced, and suggestions for improvement. No pre-simulation (baseline) test was administered, focusing this analysis on post-experience perceptions.

## Data Analysis

### *Quantitative*

Survey responses were exported to Excel for analysis. For each Likert item, we calculated descriptive statistics: mean, standard deviation, and the percentage of students who agreed or strongly agreed (i.e., responses 4 or 5). Table 1 (below) presents these results. We also visualized the data with a bar chart of percent agreement by item to illustrate key patterns.

### *Qualitative*

We applied inductive thematic analysis (Braun & Clarke, 2006) to the open-ended responses. The text were corded to identify recurring themes. The main themes that emerged were Realism and Integration, Active Engagement, Teamwork and Collaboration, and Cognitive Challenge. Representative quotes for each theme are provided in Table 2.

# Results

## Quantitative Outcomes

Table 1 displays survey results. A clear majority of students reported positive effects of the simulation. For instance, Conceptual Understanding and Theory Application both received mean ratings above 4.0, with ~ 82% of students agreeing that the simulation improved their understanding and application of course concepts. Even higher were Decision-Making and Critical Thinking: roughly 87% and 86% of students agreed they felt stronger skills in these areas, with means ~ 4.2. Engagement was exceptionally high: about 88% of respondents said the simulation was more engaging than traditional lectures (Mean ≈ 4.3). Similarly, ~80% agreed the simulation enhanced Collaboration and teamwork skills (Mean ≈ 4.1). No item scored below the neutral midpoint; all means were between 4.1 and 4.4 (SDs around 0.5 – 0.6), indicating consistently favorable perceptions.

**Table 1**
*Descriptive Statistics for Post-simulation Survey (n = 50)*

| Survey Item | % Agree/ Strongly Agree | Mean (SD) |
|---|---|---|
| The simulation improved my understanding of course concepts | 82% | 4.2 (0.6) |
| The simulation helped me apply theoretical knowledge | 80% | 4.1 (0.6) |
| The simulation enhanced my strategic decision-making skills | 87% | 4.3 (0.5) |
| The simulation improved my critical thinking skills | 85% | 4.2 (0.5) |
| The simulation was more engaging than traditional lectures | 88% | 4.4 (0.5) |
| I was actively involved in decisions during the simulation | 86% | 4.3 (0.5) |
| The simulation encouraged collaboration with classmates | 80% | 4.1 (0.6) |
| I developed better problem-solving skills in the simulation | 83% | 4.2 (0.6) |
| The simulation gave me practical insights into management | 84% | 4.2 (0.5) |
| Overall, the simulation was an effective learning tool | 85% | 4.3 (0.5) |

## Qualitative Insights

The open-ended responses corroborate the survey data. Four main themes emerged:

1. **Realism and Integration:** Many students praised how the simulation integrated multiple business functions and made theory "come alive." For example, one student noted: *"I finally saw how marketing, finance, and operations interact; making decisions in a live scenario helped me apply what I learned."* Another said it was valuable to "translate strategic plans into action," highlighting the realistic, competitive context of the game.
2. **Active Engagement:** Students described the simulation as fun and engaging. Quotes included comments like *"We learned by doing rather than just listening, which kept me interested and motivated"* and *"Making mistakes in the simulation was the best way to learn without real consequences."* This aligns with literature that simulation "stimulate[s] the enjoyment, motivation and engagement" of learners (Clarke, 2010).
3. **Teamwork and Collaboration:** Several responses emphasized teamwork. One student wrote: *"Working in a team under time pressure taught me communication and coordination skills."* Another remarked on peer learning: *"We debated and negotiated every decision; it built my confidence in group problem-solving."* About 80% of survey respondents agreed the simulation encouraged collaboration, matching these qualitative remarks.
4. **Decision Complexity and Challenge:** Students also identified challenges, such as decision complexity and time pressure. One comment was: *"It was tough to balance all variables at once, which made me think more critically."* This reflects the authentic complexity of a simulated environment.

Representative quotes are summarized in Table 2.

**Table 2**
*Qualitative Themes and Exemplary Quotes*

| Theme | Representative Quote (Student Reflection) |
|---|---|
| Realism & Integration | "The simulation linked marketing, finance, and operations. I saw how all functions interact in a real market." (Student 17) |
| Active Engagement | "Playing the simulation was much more engaging than lectures. You learn by making decisions and seeing results." (Student 8) |
| Teamwork & Collaboration | "I learned to communicate and trust my team; we had to coordinate strategies and learn from each other." (Student 22) |
| Decision Complexity | "Making quick strategic choices under pressure was challenging but realistic—it forced me to think critically." (Student 31) |

## Discussion

The findings align with prior scholarship on experiential learning and simulations. Consistent with Kolb's ELT, students reported that *active experimentation* in the simulation deepened their understanding and application of concepts. The high engagement ratings mirror Clarke's (2010) assertion that simulations stimulate motivation and allow cooperative "testing [of] hypotheses" in realistic scenarios. The strong endorsements of decision-making and critical thinking skills are also expected: business simulations intentionally create dynamic, ill-structured problems, which helps develop intuition and analysis as Clarke noted, and as Velez & Alonso's review found broadly across studies (Velez & Alonso, 2025). One student's comment about balancing multiple variables reflects how simulations compel higher-order thinking in context.

Importantly, students valued the multidisciplinary integration of the simulation. This was reflected in quotes about linking marketing, finance, and operations, and confirms that simulations allow "managing a simulated firm." Many noted that the hands-on experience made theory feel relevant—a benefit often cited by researchers (Garrison & Kanuka, 2004).

However, the study's limitations warrant cautious interpretation. All findings are based on self-reported perceptions, not objective performance measures. As reviewer feedback noted, conceptual understanding and problem-solving are best measured with tests, not just survey ratings. We acknowledge this limitation: the high percentages may partly reflect response bias or novelty effects. For example, students might overstate learning gains due to enthusiasm (or to please the instructor). There was no pre-test or control group for comparison. Future research should include objective assessments (exam scores or pre/post tests) to validate these results, and possibly a pre-activity survey to measure changes directly.

Despite these caveats, the qualitative evidence reinforces the quantitative trends. No student gave a strongly negative assessment; the constructive criticisms were about challenges of coordination and complexity, not a lack of learning. This suggests that even with self-report data, the simulation was broadly beneficial. The results imply that properly designed simulations (with clear objectives and support) can meaningfully enhance management education.

## Implications

For instructors, these findings underscore the value of integrating simulations into curricula. The high engagement suggests that simulations can make learning more active, echoing calls for "flipped" and hands-on pedagogies in higher education. The themes highlight areas to emphasize: ensure the simulation covers key functional areas of the course, facilitate team reflection to deepen learning, and be mindful of framing decisions to manage complexity. Instructors should also explicitly connect simulation tasks to theoretical frameworks in debriefing sessions, to help students articulate the links they might not notice on their own.

For researchers, the study demonstrates a straightforward way to assess classroom innovations: combining Likert surveys with thematic coding of comments. The use of both quantitative and qualitative data provides a richer picture than either alone, as recommended by mixed-methods education research literature. This approach could be applied to study other pedagogical technologies (e.g., serious games, virtual labs).

## Conclusion

This study provides evidence that classroom simulation activities can promote experiential, collaborative learning in management education. Based on a survey of MBA students, we found that a large majority perceived gains in conceptual understanding, theory application, and decision-making skills. Engagement and teamwork were also rated highly. These results align with experiential learning theory and prior research that positions simulations as powerful learning tools. We note that these conclusions rely on self-reported data, and objective assessments would strengthen future studies. Nonetheless, the findings affirm that when well-integrated into the curriculum, business simulations help students "learn by doing" and bring abstract concepts to life. As Salas et al. (2009) argue, simulation-based training has many advantages in management education, and our study shows practical evidence of those benefits in an actual classroom setting.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that AI-assisted technologies were used only to improve language, formatting, and document organization. The study design, analysis, interpretation, and conclusions remain the responsibility of the author.

## References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Clarke, E. (2010). Learning outcomes from business simulation exercises: challenges for implementation of learning technologies. *Education + Training*, *51*(5-6), 448–459, https://doi.org/10.1108/00400910910987246

Garisson, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education, 7*(2), 95–105. https://doi.org/10.1016/j.iheduc.2004.02.001

Keys, W., & Wolfe, J. (1997). *The role of management games and simulations in education and research*. *Simulation & Gaming, 28*(3), 288–303.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.

Kuczera, K. (2021). Experiential learning in simulated conditions. *Procedia Computer Science, 192*, 4186–4193. https://doi.org/10.1016/j.procs.2021.09.194

Salas, E., Wildman, J. L., & Piccolo, R. F. (2009). Using simulation-based training to enhance management education. *Academy of Management Learning & Education, 8*(4), 559–573.

Velez, A., & Alonso, R. (2025). Business simulation games for the development of decision making: Systematic review. *Education Sciences, 15*(2), 168. https://doi.org/10.3390/educsci15020168

# Industry-Integrated Curriculum Design for Employability in Marketing Education: The BRIDGE Process Framework

Ying Wu, University of Dundee, United Kingdom
Kirsty Beaumont, University of Dundee, United Kingdom
Ruth O'Riordan, University of Dundee, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Graduate employability is a strategic priority in business and marketing education, yet a persistent gap remains between conventional curricula and the audience-facing work of contemporary practice, and the field lacks an empirically derived framework specifying how that gap might be addressed through curriculum design rather than extra-curricular provision. This study examines how industry-integrated curriculum design creates structured opportunities for students to enact and evidence employability competencies. It adopts a longitudinal, embedded, qualitatively driven mixed-method case study of one undergraduate and one postgraduate marketing module in a UK business school across four cohort iterations. The corpus comprises student evaluation comments, in-lecture feedback and co-design responses, curriculum documents, reflective teaching records and institutional assessment records, analysed through abductive thematic analysis, cross-case comparison, temporal sequence analysis and a descriptive quantitative strand integrated by joint display. The postgraduate module affords a natural contrast between cohorts taught with and without a live client, holding assessment format constant. Industry integration was associated with higher reported relevance and created opportunities to enact digital marketing strategy, analytical reasoning, collaborative problem solving and professional communication, conditional on client-task fit, workload calibration and scaffolding. Postgraduate marks were descriptively higher under the live-client condition, though intact-cohort confounds preclude causal inference, while documented undergraduate curriculum improvements were not accompanied by higher marks. The study develops the BRIDGE framework as a process model of inputs, learning mechanisms, a proximal outcome and an adaptive mechanism, specified through three propositions and four boundary conditions, and advances curriculum evaluation as a dual-tempo adaptive mechanism.

*Keywords:* employability, work-integrated learning, authentic assessment, marketing education, curriculum design

iafor
The International Academic Forum
www.iafor.org

## Introduction

Graduate employability sits at the centre of strategic planning in UK business schools, driven by employer expectations that graduates combine disciplinary knowledge with applied competencies (Bennett, 2019; Clarke, 2018; Tight, 2023). Succi and Canovi (2020) report a widening perception gap: employers rate applied skills more highly than students recognise they need. This matters acutely in marketing, whose constructs are validated through application to real markets (Grewal et al., 2025). A curriculum teaching those constructs only through simulated cases risks producing graduates who can describe strategic decisions but have never made one for a real audience with reputational consequences.

The gap is structural rather than incidental (Jackson, 2016; Whittaker, 2016): it reflects how curricula are assessed, through summative, staff-graded, low-stakes simulation, rather than deficiency of content (Coates et al., 2016). Embedding authentic tasks into formal assessment rather than adding employability extra-curricularly (Yorke & Knight, 2006) remains inconsistently applied (Jackson et al., 2023), and client-based projects rarely scale beyond instructor enthusiasm (Shanahan et al., 2021).

Three gaps motivate this paper. First, employability frameworks are largely conceptual or diagnostic, describing what employability consists of without specifying a testable design procedure (Eimer & Bohndick, 2023). Second, evidence spanning undergraduate and postgraduate curricula within one institutional case is scarce, so it is unclear whether integration mechanisms operate consistently across levels. Third, comparison with conventional simulated assessment is rarely made explicit. The study therefore asks:

- **RQ1:** How does industry-integrated curriculum design create structured opportunities for students to enact and evidence employability competencies?
- **RQ2:** How do the mechanisms and boundary conditions of industry-integrated curriculum design differ from those of conventional simulated assessment?

The aim pursued through these questions is to elaborate abductively a transferable process framework, extending related prior work on industry-collaborative and gamified assessment (Wu & Stewart, 2024; Wu et al., 2025) from single mechanisms to an integrated model. The contribution is BRIDGE, specified through three propositions and four boundary conditions, whose principal advance is curriculum evaluation as a dual-tempo adaptive mechanism. The study does not measure employability gain, and distinguishes throughout designed opportunity and demonstrated enactment from measured improvement.

## Theoretical Background

The grounding was developed through an integrative narrative review establishing sensitising concepts for abductive analysis, not a systematic review; no screening protocol or review trail of the kind required by systematic methodology (Tranfield et al., 2003) was maintained, and the review is not presented as reproducible. The resulting synthesis is summarised in Tables 1 and 2, mapping each strand to its core contribution and to the use made of it here.

### Work-Integrated and Authentic Learning

Work-integrated learning denotes approaches purposefully integrating theory with workplace-relevant practice (Billett, 2009), requiring capacity to navigate ambiguity rather than execute defined procedures (Oliver, 2015). Cooper et al. (2010) and Patrick et al. (2008) stress that

quality depends less on a partner's presence than on how deliberately the partnership is designed into curriculum, a claim examined abductively here. Authentic learning theory explains why live-client work should differ from simulation: tasks are situated in real contexts with genuine stakes (Herrington & Oliver, 2000), with ill-defined problems, a polished complete product and an audience beyond the assessor (Ashford-Rowe et al., 2014). Kolb (1984) underpins the iterative logic and Schon (1983) reflective adaptation. Industry feedback must be formalised within assessment (Richardson & Henschke, 2010) and feedback designed rather than appended (Boud & Molloy, 2013). Smith (2012), Rowe and Zegwaard (2017), Jackson (2015) and Ferns et al. (2016) converge on design quality, not client presence, as decisive; Bandaranaike and Willison (2015) and Trede (2012) extend this to affective readiness and professional identity.

## Employability Capital

Harvey (2001) and Knight and Yorke (2003) established employability as a curricular rather than outcome property. Bridgstock (2009) identifies career management skills as overlooked; Tomlinson (2017) proposes five interacting capitals, and Kalfa and Taksa (2015) show cultural capital shapes how employability is recognised in business education. Hinchliffe and Jolly (2011) and Holmes (2013) debate whether employability is possession, position or process, which Jackson (2016) partly resolves through pre-professional identity, while Jackson and Wilton (2017) show self-management matters as much as experience. Jackson et al. (2023) address assessing work-integrated learning absent external accreditation. Kinash et al. (2016) show that stakeholders disagree about which employability strategies work, and Helyer and Lee (2014) situate the debate within policy expectations. Eimer and Bohndick (2023) find most models remain diagnostic rather than translated into design procedure, the gap BRIDGE addresses.

## Interdisciplinarity and Mode 2 Knowledge

Interdisciplinary learning integrates concepts from two or more disciplines to address one problem (Laasch et al., 2023); transdisciplinary learning involves non-academic stakeholders as co-producers (Janssens et al., 2022). Gibbons et al. (1994) and Nowotny et al. (2001) distinguish disciplinary knowledge from knowledge produced in the context of application, a distinction this paper uses to bound rather than inflate its transdisciplinary claim. University-industry engagement generates more durable value when relational rather than transactional (Ankrah & Al-Tabbaa, 2015; Perkmann et al., 2013). Curriculum co-creation with students (Bovill, 2020; Healey et al., 2014; Zainudden et al., 2022) provides a template extended here to include the industry partner.

**Table 1**

*Literature Synthesis by Thematic Strand (Theoretical Foundations)*

| Strand | Key sources | Core contribution | Use in this study |
|---|---|---|---|
| Employability as institutional priority | Harvey (2001); Knight and Yorke (2003); Clarke (2018); Tight (2023); Bennett (2019); Succi and Canovi (2020); Grewal et al. (2025) | Reframes employability as capital and context rather than a fixed attribute list; evidences accelerating skills demand in marketing | Establishes urgency and disciplinary specificity of the gap |
| Situated and experiential learning | Lave and Wenger (1991); Brown et al. (1989); Wenger (1998); Kolb (1984); Schon (1983) | Locates learning in participation and context rather than transmission | Underpins why an external audience changes the learning task |
| Work-integrated learning quality | Billett (2009); Cooper et al. (2010); Patrick et al. (2008); Smith (2012); Rowe and Zegwaard (2017); Jackson (2015); Ferns et al. (2016); Oliver (2015) | Specifies quality dimensions and partner-capacity requirements; identifies design, not partner presence, as decisive | Directly motivates P1 and the boundary conditions |
| Authentic assessment and feedback design | Herrington and Oliver (2000); Ashford-Rowe et al. (2014); Villarroel et al. (2018); Sadler (1989); Nicol and Macfarlane-Dick (2006); Boud and Molloy (2013); Carless and Boud (2018); Richardson and Henschke (2010) | Distinguishes authentic from simulated assessment; establishes feedback as designed rather than appended | Specifies the D component and motivates P2 |

**Table 2**

*Literature Synthesis by Thematic Strand (Application and Method)*

| Strand | Key sources | Core contribution | Use in this study |
|---|---|---|---|
| Employability capital and identity | Bridgstock (2009); Tomlinson (2017); Jackson (2016); Jackson and Wilton (2017); Holmes (2013); Kinash et al. (2016); Jackson et al. (2023); Eimer and Bohndick (2023) | Conceptualises employability as multidimensional capital and identity; identifies the assessment gap in non-accredited settings | Diagnoses the gap the I component addresses |
| Interdisciplinarity and Mode 2 knowledge | Gibbons et al. (1994); Nowotny et al. (2001); Laasch et al. (2023); Janssens et al. (2022); Leal Filho et al. (2019) | Distinguishes disciplinary from context-of-application knowledge production and co-production with non-academic actors | Provides the criterion against which the transdisciplinary claim is qualified |
| University-industry engagement and co-creation | Perkmann et al. (2013); Ankrah and Al-Tabbaa (2015); Healey et al. (2014); Bovill (2020); Zainudden et al. (2022); Shanahan et al. (2021) | Shows relational engagement outperforms transactional; establishes co-creation with students as design practice | Frames the R component and the three-way design relationship |
| Qualitative theory building | Eisenhardt (1989); Gioia et al. (2013); Corley and Gioia (2011); Stake (1995); Yin (2018); Braun and Clarke (2019); Lincoln and Guba (1985); Tracy (2010); Nowell et al. (2017); Miles et al. (2014) | Specifies procedures for abductive framework development and criteria for evaluating contribution | Provides the methodological warrant for the study |

# Methodology

## Design

The study adopts a longitudinal, embedded, qualitatively driven convergent mixed-method case study (Eisenhardt, 1989; Eisenhardt & Graebner, 2007; Stake, 1995; Yin, 2018), informed by constructive alignment (Biggs & Tang, 2011). It is *embedded* (two curricular subcases), *longitudinal* (four cohort iterations with documented change) and *qualitatively driven*: the qualitative strand carries analytic weight and generates the framework, while assessment records contextualise, corroborate or challenge it. The quantitative strand is descriptive; no inferential test is performed and integration proceeds by joint display.

## Scope

The study examines how arrangements create opportunities for enactment and how they were adapted; it does not measure employability gain. The postgraduate module provides a natural contrast between cohorts taught without and with a live client, holding format constant, but these are intact groups differing in year, size and composition.

## Corpus

Table 3 sets out the corpus, the enrolled population and contributor counts for each source, together with the principal evidential limitation attaching to it. **Cases.** BU31013 (undergraduate) assesses a 2.5-minute SDG-aligned podcast with show notes (30%) and a SOSTAC campaign plan (70%), drawing on sustainability marketing typologies (Kemper & Ballantine, 2019) and SDG framing (Leal Filho et al., 2019). Its live client was Intelligent Growth Solutions (IGS), a Dundee vertical-farming technology company selling into business and public-sector markets, in 2024/25, and Campy Growers, a consumer-facing community growing initiative, in 2025/26. BU52039 (postgraduate) assesses a group brand-extension pitch (30%) and a 3,000-word individual report (70%), applying brand-extension theory (Aaker & Keller, 1990). In 2024/25 students selected their own brand with no external stakeholder; in 2025/26 the client was Lifelong Learning Dundee. Review forms confirm structure, weighting and word counts were unchanged. The modules were selected as theoretically informative: the undergraduate module embedded a live-client brief in both observed iterations, whereas the postgraduate module did so only in 2025/26, and the two differ in level, client context, format and client-facing interaction.

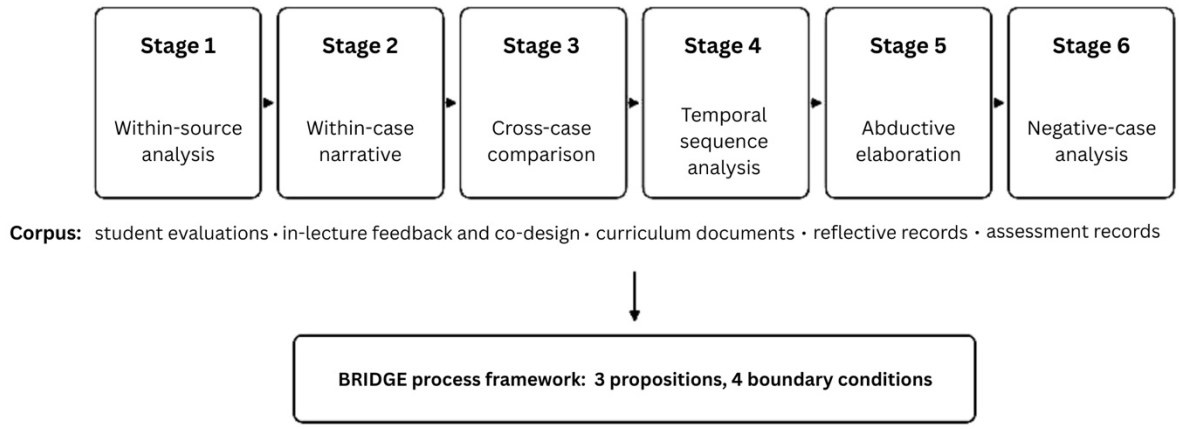
**Table 3**
*Corpus Inventory and Evidential Limitations*

| Source | Context | Enrolled | Contributors | Principal evidential limitation |
|---|---|---|---|---|
| Module evaluation, open text | UG 2024/25 | 56 | 12 and 10 per prompt | Not purpose-built for employability; self-selecting |
| Module evaluation, open text | UG 2025/26 | 59 | 7 and 7 per prompt | Low response rate; bias unquantifiable |
| Module evaluation, open text | PG 2025/26 | 20 | 4 and 4 per prompt | Very small base; provisional only |
| Weekly in-lecture feedback | UG, full semester | Attendance varied | 13 to 31 per session | Public setting; social desirability likely |
| In-lecture quiz records | Both modules | Attendance varied | 7 to 14 per item | Evidences provision, not gain |
| Co-design responses | PG 2025/26 | 20 | 4 | Four respondents; indicative |
| Curriculum documents | Both modules | n/a | n/a | Design intent, not enactment |
| Reflective teaching records | Four iterations | Author team | 3 authors | Authored by stakeholders in the account |
| Assessment records | All iterations | 191 | 182 marked | Intact cohorts; no baseline |

## Analysis

Six stages were used (Creswell & Poth, 2018; Miles et al., 2014; Patton, 2015): within-source analysis, recognising documents and reflective records as researcher-created artefacts rather than informant accounts; within-case narrative per subcase, informed by structured reflective cycles (Gibbs, 1988); cross-case comparison distinguishing recurring from contingent mechanisms; temporal sequence analysis from signal to action; abductive elaboration against the theoretical background; and negative-case analysis specifying boundary conditions. Figure 1 sets out the sequence and its relation to the corpus and to the resulting framework.

**Figure 1**
*Six-Stage Abductive Analytical Procedure*




*Note.* Sources are analysed separately before synthesis; the framework is elaborated at Stage 5 and bounded at Stage 6.

Consistent with reflexive thematic analysis (Braun & Clarke, 2019), coding was interpretive rather than a reliability exercise: the lead author coded iteratively, co-authors acted as critical interpretive partners challenging themes and contribution claims, and divergences were retained where meaningful. No inter-rater reliability claim is made. Because the corpus mixes informant accounts with design artefacts, Table 4 is described as Gioia-inspired (Gioia et al., 2013) rather than as an application of that methodology.

Three quantities are distinguished: enrolment, assessment completion and qualitative contribution. Gradebook records are authoritative for cohort size. Across live-client iterations 128 students were marked on the principal assessment; the qualitative sample is far smaller, and findings should be weighed against contributor counts.

**Trustworthiness**

Six strategies were used (Lincoln & Guba, 1985; Nowell et al., 2017; Tracy, 2010): source triangulation with attribution; cross-case testing of every component (Table 4); temporal traceability; negative-case analysis; a reflexive audit trail; and peer challenge by an author not responsible for the module concerned. The authors are insider researchers, a limitation mitigated but not removed by reporting critical evidence in full. **Ethics.** Approved by the School of Business Research Ethics Committee, University of Dundee (UoD-SoB-STAFF-2025-16). In-lecture participation was voluntary and anonymous at collection and carried no assessment consequence. No student names, identifiers or platform display names were accessed or reported.

Table 4 reports the source supporting each item, so that student-supported claims are distinguished from claims resting on the authors' reflective records. It also retains evidence that resolves into boundary conditions rather than confirming the framework: a structure in which every item supported the framework would risk suggesting an analysis shaped by its conclusion.

**Table 4**
*Evidence Structure With Cross-Case Status and Evidential Strength*

| Illustrative evidence | Source | Interpretive theme | Dimension | Cross-case status | Strength |
|---|---|---|---|---|---|
| "Learning about Campy Growers... helped me understand how local initiatives can create meaningful impact" | UG evaluation | Authentic framing distinguishes brief from simulation | B: Brief | Both subcases | Moderate |
| Client reviewed outputs; pitch delivered to external stakeholder | Documents; reflective record | Sustained engagement with external stakeholder | R: Real Client | Both subcases | Strong (documented) |
| Rubric criteria naming digital strategy, analytical reasoning, communication | Briefs and rubrics | Explicit mapping of criteria to competency domains | I: Integrated Competencies | Both subcases | Strong (documented) |
| "Have seminar activity feed into the assignments"; exemplar and earlier-content requests | PG co-design; UG weekly feedback | Demand for anticipatory scaffolding aligned to summative task | D: Developmental Feedback | Both subcases | Moderate |
| Podcast, pitch deck and report produced as standalone artefacts | Briefs and documents | Artefacts designed for external legibility | G: Graduate-Ready Evidence | Both subcases | Design intent only |
| Client changed; exemplars released earlier; fortnightly quiz cycle adjusted | Reflective record; documents | Recalibration at two tempos | E: Evaluation for Enhancement | Slow loop both; fast loop reported only | Mixed |
| "Maybe a B2C company would be better next year"; "most time-consuming... report I have done" | UG evaluation | Client-context misalignment; workload disproportion | Boundary condition | UG, echoed in PG | Strong (recurrent) |
| "Using the client was quite hard as they have constrains/ limited budget" | PG co-design | Client capacity constrains achievable scope | Boundary condition | PG only | Indicative |

## Findings

### Reported Relevance

Students attributed engagement to the live-client format rather than delivery alone: “Learning about Campy Growers and their work in Dundee helped me understand how local initiatives can create meaningful social and environmental impact” (UG 2025/26). A postgraduate described the relationship as reframing theory: “it is good because it's very practical and transferable.” This accords with authentic learning theory (Herrington & Oliver, 2000) and evidence that career-oriented learning shapes perceived employability (Ho et al., 2023; Huang, 2025).

### Opportunities to Enact Competencies

The design created structured opportunities to enact four domains: digital marketing strategy, through SOSTAC planning against live audience data; analytical reasoning, through competitor mapping on a real organisation; collaborative problem solving, through a group pitch under real constraints; and professional communication, through client-facing formats. The postgraduate pitch applied brand-extension theory to a live client decision rather than textbook examples. The claim concerns designed opportunity and observable enactment, not measured gain.

## Professional Register

Client-facing formats required students to address a lay or client audience rather than an academic marker, a competency rarely assessed directly in essay-based curricula.

## Tensions

Undergraduates in 2024/25 found the business-to-business client hard to work with: “it is a bit harder to apply to a B2B organisation... maybe a B2C company would be better next year.” Workload criticism was sharp: “the most time-consuming and difficult 2,000 report I have done.” A postgraduate noted capacity limits: “using the client was quite hard as they have constrains/ limited budget.” These recur across cohorts, indicating a genuine tension between authenticity, which inherits a client's real complexity, and a single module's bounded workload. This is the mechanism Sadler (1989) and Carless and Boud (2018) would predict: without the means to interpret and act on standards in advance, ill-defined tasks convert into effort rather than into learning.

## Assessment Records

Marks use the University's 23-point scale; percentage conversion is meaningless here and platform percentage bands are disregarded; Table 5 reports the distributions. One matter requires disclosure before the figures are read. Three distributions contain a recorded mark of zero. These zeros record non-submission rather than an assessed but failing attempt; no academic-offence outcome is involved. Because a non-submission does not represent assessed performance, zeros are reported as recorded for completeness and a sensitivity check excluding the single lowest mark is reported alongside, since their inclusion materially affects dispersion. Claims about the floor of the distribution are not advanced, since the second-lowest mark cannot be recovered from summary statistics.

**Table 5**

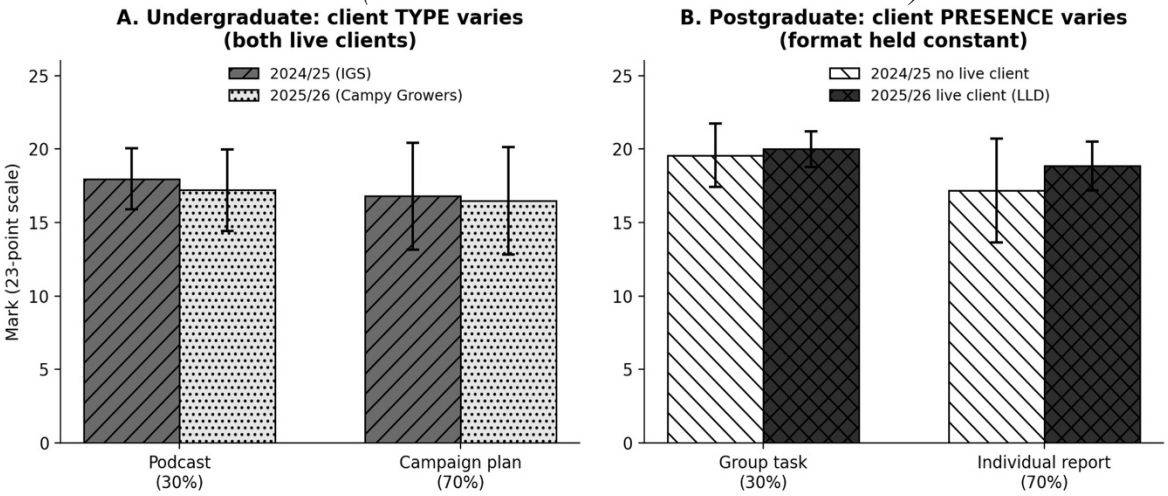
*Mark Distributions by Client Condition (23-Point Scale)*

| Iteration | Client condition | Assessment | n | Mean | SD | Min |
|---|---|---|---|---|---|---|
| BU31013 2024/25 | Live (IGS, B2B) | Podcast (30%) | 50 | 17.98 | 2.08 | 14 |
| BU31013 2024/25 | Live (IGS, B2B) | Campaign plan (70%) | 52 | 16.80 | 3.63 | 0 |
| BU31013 2025/26 | Live (Campy Growers) | Podcast (30%) | 54 | 17.22 | 2.77 | 5 |
| BU31013 2025/26 | Live (Campy Growers) | Campaign plan (70%) | 56 | 16.48 | 3.65 | 0 |
| BU52039 2024/25 | **No live client** | Group presentation (30%) | 56 | 19.58 | 2.16 | 11 |
| BU52039 2024/25 | **No live client** | Individual report (70%) | 54 | 17.18 | 3.53 | 0 |
| BU52039 2025/26 | **Live (LLD)** | Group pitch (30%) | 16 | 20.00 | 1.22 | 18 |
| BU52039 2025/26 | **Live (LLD)** | Individual report (70%) | 20 | 18.85 | 1.67 | 15 |

**Table 6**
*Two Natural Contrasts: Analytic Structure and Inference*

| Contrast | What varies | Held constant | Direction | Descriptive size | Analytic inference | Principal confound |
|---|---|---|---|---|---|---|
| UG 2024/25 to 2025/26 | Client *type* (B2B to consumer-facing) | Client presence; format; level; team | Podcast -0.76; plan -0.32 | Negligible; within dispersion | Client type alone did not raise attainment though feedback tenor improved: curriculum improvement and performance gain are separable | Scaffolding changed simultaneously |
| PG 2024/25 to 2025/26 | Client *presence* (none to live) | Format; weighting; word counts; level | Report +1.67; group +0.42 | d = 0.53 and 0.21; d = 0.56 excluding one non-submission | Difference appears where presence varies, not type; consistent with P1 | Cohort size 56 to 20; year; composition |

**Figure 2**
*The Two Natural Contrasts (Error Bars Show One Standard Deviation)*




*Note.* Panel A varies client type with presence held constant; Panel B varies client presence with assessment format held constant. Bars are distinguished by fill pattern as well as tone.

The design contains two analytically distinct comparisons, displayed in Figure 2 and set out analytically in Table 6. Two comparisons are distinguished. The undergraduate comparison varies client **type** while holding client presence constant; the postgraduate comparison varies client **presence** while holding assessment format constant, and is the closer approximation to a comparison condition.

The contrasts are more informative together because they point in opposite directions. A reading in which any live client raises attainment is inconsistent with the undergraduate row; a reading in which presence is immaterial is inconsistent with the postgraduate row. The pattern is at least consistent with the possibility that what a live client supplies is a set of constraints, an external audience, a real problem and a recipient of the output, which the assessment then either exploits or fails to exploit; motivation was not measured and no claim about it is made here. On that reading, changing which organisation supplies those constraints should matter less than whether they are present at all, which is what Table 6 shows. Effect sizes are descriptive, not tests. Because the no-client distribution contains a recorded zero arising from non-submission rather than from assessed performance, a sensitivity check excluding the single lowest mark was run: the no-client mean becomes 17.50 (SD = 2.63, n = 53) and d = 0.56. The

direction and approximate magnitude are robust to that exclusion, but any claim of narrowed dispersion is not, since much of the gap between SD = 3.53 and SD = 1.67 rests on a single value that does not represent assessed performance. Claims about dispersion and about the floor of the distribution are therefore withdrawn. The postgraduate confound remains decisive: a difference of this size between intact cohorts differing threefold in size is compatible with cohort variation alone. The finding is convergent, not confirmatory.

### Adaptation

Both modules ran a fortnightly formative quiz cycle alongside session-level feedback prompts. Table 7 traces each documented sequence from feedback signal to curriculum action, and states what each sequence does and does not license.

**Table 7**
*Feedback-to-Action Sequences and What Each Licenses*

| Timing | Signal | Action | Evidential status | What this licenses |
|---|---|---|---|---|
| End UG 2024/25 | B2B client hard to apply concepts to | Client changed to consumer-facing organisation | **Completed**; curriculum documents | Slow loop, independent of author testimony |
| End UG 2024/25 | Structure shared close to deadline | Structure and exemplars released earlier | **Completed**; curriculum documents | Slow loop; evidences scaffolding boundary |
| Within semester, both | Content requested earlier | Sequencing adjusted in later sessions | **Reported**; reflective record only | Actor, decision, implementation named; not verifiable |
| Within semester, UG | Worked example requested | Exemplar provided; formalised next iteration | **Reported** then **completed** | Spans both tempos; strongest fast-loop evidence |
| Within semester, both | More formative quizzing requested | Fortnightly quiz provision adjusted | **Reported**; records evidence provision only | Provision established, not causal sequence |
| PG 2025/26 | Weighting and client capacity concerns | Recorded for next iteration | **Need surfaced**; action pending | Loop instrumented; no closure observed |

The loops are evidenced differently. Between-cohort change is documented and independently verifiable; within-semester change rests on the team's contemporaneous account, which names actor, decision and implementation but originates with authors who have a stake in it and lacks a dated log. That asymmetry bounds the dual-tempo claim: the fast loop is demonstrably instrumented and generates actionable signal, while its closure is attested rather than documented.

## The BRIDGE Process Framework

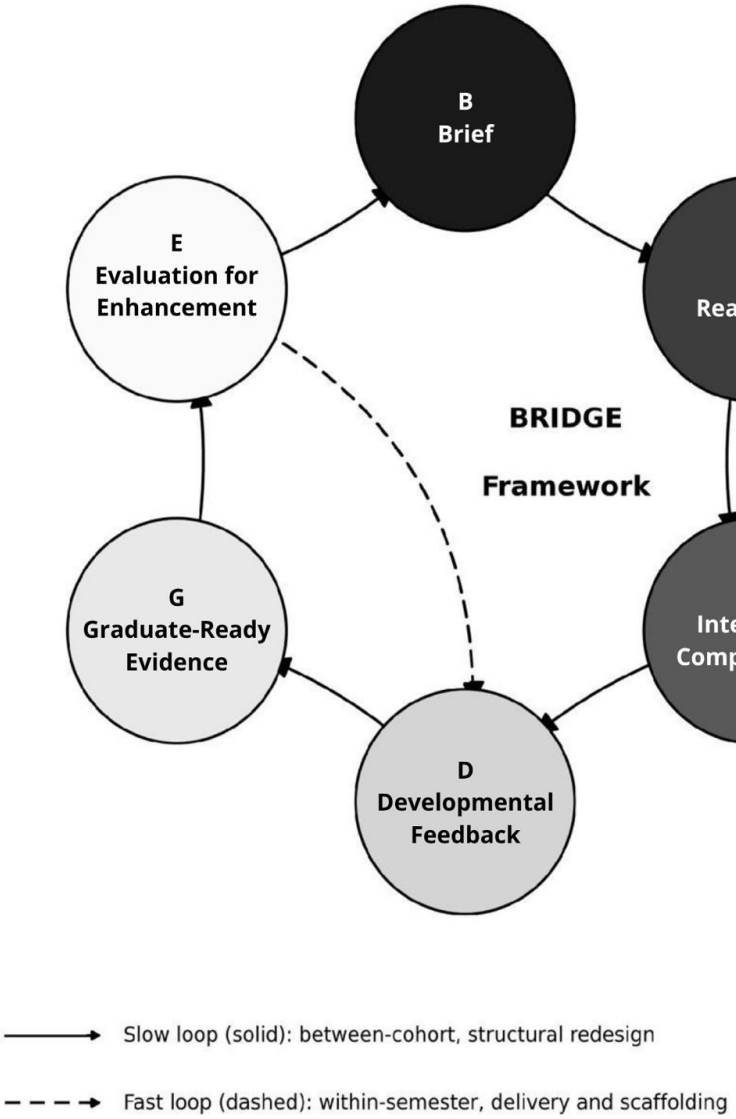
A framework whose components sit at different conceptual levels invites the objection that it is an acronym rather than a model. That objection is answered by treating the heterogeneity as structure.

**Table 8**
*BRIDGE Components by Conceptual Role, With Falsification Conditions*

| Role | Components | Function in the process | Evidenced in | Would be falsified by |
|---|---|---|---|---|
| Inputs | B: Brief; R: Real Client | Establish authentic problem and external audience with genuine stakes | Both subcases (documented) | Authentic briefs yielding no difference in relevance or enactment |
| Learning mechanisms | I: Integrated Competencies; D: Developmental Feedback | Make competency demands explicit; render authentic uncertainty productive rather than overwhelming | Both subcases (documented and reported) | Criteria and scaffolding not altering how workload is experienced |
| Proximal outcome | G: Graduate-Ready Evidence | Yield artefacts legible to external professional audiences | Design intent only; downstream use unmeasured | Artefacts unusable beyond assessment |
| Adaptive mechanism | E: Evaluation for Enhancement | Recalibrate delivery within semester (fast loop) and structural design between cohorts (slow loop) | Slow loop documented; fast loop reported | Curricula adapting equally well without a fast channel |

The ordering is temporal and mechanistic: brief and client create conditions for enactment, made explicit by integrated criteria and survivable by developmental feedback, producing legible artefacts, which with evaluation feed recalibration of delivery within semester and of the brief between cohorts (Figure 3).

**Figure 3**
*The BRIDGE Process Framework as a Dual-Tempo Cycle*




*Note.* Fast loop (dashed): Evaluation to Developmental Feedback within semester. Slow loop: to Brief between cohorts.

## Propositions

Industry integration does not strengthen employability learning simply because a client is present; value arises where authentic stakes combine with explicit competency integration, scaffolding, legible evidence and recurring evaluation.

- **P1. Authenticity must be structured to be educative.** A live brief establishes relevance but not learning; educative value depends on whether criteria require integration of disciplinary and professional competencies. Table 6 is consistent: difference appears where presence varies, not where type varies.
- **P2. Developmental feedback conditions whether authentic uncertainty is experienced as productive challenge or disproportionate overload.** Evidenced by co-occurrence of strong endorsement of formative provision with sharp workload criticism where provision arrived late.

- **P3. Employability-oriented curricula require adaptive evaluation at two tempos.** A fast within-semester loop recalibrates delivery for the cohort enrolled; a slow between-cohort loop revises structural choices such as client selection. Running only the slow loop improves conditions for future students at present ones' expense.

### Boundary Conditions

(1) **Client-task fit:** where a client's market position does not map onto taught concepts, authenticity costs accessibility.
(2) **Client capacity:** where budget or operational limits constrain implementable recommendations, the task feels artificially bounded.
(3) **Workload proportionality:** where effort is not reflected in weighting, dissatisfaction follows irrespective of pedagogical merit.
(4) **Scaffolding lead time:** where structure arrives close to submission, the fast loop cannot operate for the cohort needing it.

## Discussion

**Table 9**
*Structural Comparison With Ideal-Typical Simulated Design, and the Mechanism at Stake*

| Dimension | Ideal-typical simulated design | Industry-integrated design | Mechanism at stake |
|---|---|---|---|
| Assessment stakes | Simulated case, no external audience | Real client brief with reputational relevance | External audience redefines adequacy |
| Feedback timing | End-point summative only | Iterative formative checkpoints | Determines if uncertainty is productive |
| Skill focus | Disciplinary knowledge tested abstractly | Knowledge plus applied competency, explicitly mapped | Criteria make demands assessable |
| Evidence of employability | Grade and transcript | Professionally oriented artefacts | Evidence becomes portable |
| Stakeholder involvement | Academic staff only | Staff and industry partner jointly | External framing resists simplification |
| Principal risk | Lower relevance; weaker transfer | Workload and accessibility unevenly distributed | Complexity cannot be calibrated |

That comparison is set out in Table 9, which names the mechanism at stake in each dimension. It isolates a design type, not a description of practice; many conventional modules include formative feedback and professional outputs. The final row is deliberate: industry integration carries a risk profile a staff-authored case does not, because complexity cannot otherwise be calibrated.

### Interdisciplinary Integration

Within one brief students integrated sustainability and SDG framing, digital analytics, strategic brand management and professional communication, four domains conventionally taught separately, meeting the criterion of cross-domain integration. The transdisciplinary claim is weaker: the design contains elements of transdisciplinary engagement where clients contributed to problem framing and responded to outputs, but clients did not frame learning outcomes, interpret evidence or shape the framework. Claiming otherwise would inflate engagement into co-production.

The joint display in Table 10 integrates the qualitative and quantitative strands, marking convergence, divergence and silence.

## Locating The Dual-Tempo Mechanism Theoretically

The claim that curriculum evaluation operates at two tempos is not merely descriptive, and it can be located within an established account of organisational learning. Argyris and Schon (1978) distinguish single-loop learning, in which error is detected and corrected within existing governing variables, from double-loop learning, in which the governing variables themselves are interrogated. The two loops observed here map onto that distinction, though imperfectly and the imperfection matters. The within-semester loop corrects delivery, sequencing and scaffolding while holding the design parameters of the module fixed, which is single-loop in form: the brief, the client and the assessment architecture are not in question. The between-cohort loop was the only one in which a governing variable was actually revised, when the client itself was changed after the business-to-business brief proved difficult to reconcile with the taught curriculum. That is closer to double-loop learning, because what was revised was not the execution of the design but a premise of it.

The mapping is imperfect because Argyris and Schon distinguish the loops by what is interrogated, whereas the present distinction is compounded by tempo: in a credit-bearing module, questioning a governing variable is normally impossible mid-delivery, so the depth of a change and its timing are confounded by institutional constraint rather than by choice. This is itself a finding worth stating. Curriculum design is a setting in which double-loop learning is structurally deferred, since the parameters most worth interrogating, client selection, assessment weighting and task scope, are precisely those that regulation and student expectation prevent altering for the cohort that surfaced the problem. The practical consequence follows directly: teaching teams should not treat responsiveness within a semester as evidence that a design is sound, because the fast loop is constitutionally incapable of testing the design premises. Conversely, a team that runs only the slow loop has no mechanism for detecting error at all until the cohort has left. The argument for instrumenting both is therefore stronger than convenience: they detect different classes of error, and neither substitutes for the other.

**Table 10**

*Joint Display: Convergence, Divergence and Silence Across Strands*

| Mechanism | Qualitative evidence | Quantitative evidence | Integrated interpretation |
|---|---|---|---|
| B and R | Engagement attributed to live-client format; B2B client harder to apply concepts to | PG mean higher with live client; UG unchanged when type changed | **Convergent but conditional.** Presence and fit separable |
| I | Comments on practical application; rubric criteria naming domains | No rubric-level data available | **Silent.** Competency-level performance unobservable here |
| D | Requests for exemplars, earlier content, more quizzing; workload criticism | PG mean higher; d = 0.53 (0.56 excluding one non-submission) | **Weakly convergent.** Cohort size equally plausible |
| G | Artefacts designed for external audiences; professional register demands | Not addressed | **Silent.** No evidence beyond assessment |
| E | Documented client change; fortnightly quiz cycle; reported within-semester adjustment | UG marks did not rise after documented changes | **Divergent.** Experience improved; attainment did not |

The display exposes silence and divergence, not only agreement. Two mechanisms are quantitatively unevidenced and marked as such rather than inferred. The final row records a

genuine disagreement, retained rather than reconciled: changes students experienced as improvements were not accompanied by higher marks. This constrains what the adaptive mechanism can claim, and supports reading BRIDGE as a specification of design mechanisms and student experience rather than an attainment intervention.

### Rival Explanations

Three alternative readings of the pattern deserve statement, since a case study that considers only its preferred interpretation offers weak grounds for belief. The first is selection: the postgraduate cohorts differ nearly threefold in size, and a smaller cohort may be differently composed in prior attainment, language background or programme mix. Nothing in the available records excludes this, and it alone could generate the observed difference. The second is maturation of the teaching team: by 2025/26 the same staff had delivered the design repeatedly, so any improvement may reflect instructor learning rather than the presence of a client. The undergraduate contrast offers weak evidence against this reading, since marks there did not improve despite the same accumulation of experience, but the two modules are not equivalent and the inference is indirect. The third is expectancy: students and markers aware that a live client was involved may have adjusted effort or judgement accordingly, and marking was not blinded to condition. This study cannot adjudicate between these explanations and the mechanism it proposes, and does not claim to. What it can do is specify what would discriminate between them, which is the function of the falsification conditions in Table 8 and of the future research agenda.

## Implications for Practice

Four implications follow, each tied to an evidenced mechanism rather than to general advocacy. First, live-client selection is a pedagogical decision, not an administrative one. The undergraduate contrast shows that authenticity was constant while student experience was not, and the boundary condition on client-task fit explains why: a client whose market position sits outside the concepts being taught imposes a translation burden that competes with the intended learning. Teams should audit prospective clients against the taught curriculum before agreeing a brief.

Second, scaffolding lead time is the variable most within a teaching team's control and the one most consistently implicated in dissatisfaction. Exemplars and assessment structure released late in the semester cannot serve the cohort that needs them, whatever their eventual quality. Where resource is limited, earlier release of existing materials is likely to yield more than the creation of new ones.

Third, the fast evaluation loop was implemented through lightweight recurring prompts in these modules and, on this evidence, generated signals that end-of-module instruments do not surface. No cost data were collected, so the burden in other settings is unknown. A recurring in-session prompt produced specific, actionable proposals on weighting, sequencing and formative provision; the postgraduate co-design prompt produced assessment-design proposals from students who had just completed the task. Teams adopting BRIDGE should treat the absence of a within-semester channel as a design gap rather than an acceptable simplification, and should record what was changed and when, which this study could not.

Fourth, client relationships should be scoped against capacity. The postgraduate observation that a client's budget constrained implementable recommendations indicates that partner

capacity shapes what students can meaningfully propose. Where capacity is limited, splitting assessment so that one task serves the client and another addresses broader disciplinary content, as students themselves proposed, preserves authenticity without narrowing the curriculum.

## Contribution, Limitations and Conclusion

### Contribution

Against Corley and Gioia's (2011) criteria, the paper claims one theoretical advance rather than many: curriculum evaluation specified as a dual-tempo adaptive mechanism, where the literature reviewed largely treats evaluation as a single retrospective activity. Its corollary is that a curriculum running only the slow loop improves conditions for future students at the expense of present ones. The literature reviewed did not identify a framework integrating work-integrated learning, authentic learning, employability capital and interdisciplinary design into one mechanism-level process framework derived abductively; this is a claim about the corpus reviewed, not the field. Practically, BRIDGE specifies mechanism rather than marketing content, is portable to other applied disciplines, and equips adopting teams with evidenced cautions rather than aspiration.

### Scope Conditions for Transfer

Because the framework specifies mechanism rather than marketing content, its transferability depends on conditions that should be stated rather than assumed. Three appear necessary on this evidence. The discipline must admit of problems that an external organisation genuinely holds, so that the brief is not a simulation with a nominal client attached; where no such problem exists, the Brief and Real Client components collapse into conventional case work. The assessment architecture must permit criteria to be written against competencies rather than content coverage alone, since without this the Integrated Competencies component has nothing to attach to. And the delivery pattern must include recurring contact of sufficient frequency to carry a fast evaluation loop; a block-taught or wholly asynchronous module may be unable to operate one, in which case the framework reduces to its slow loop and the paper's central claim does not apply. Programmes in which these conditions do not hold are not poor candidates for industry engagement, but they are poor candidates for this framework, and saying so is more useful than claiming general applicability.

### Limitations

The design cannot show employability improved, and does not claim to. The qualitative sample is modest relative to enrolment, particularly postgraduate, and self-selection is unquantifiable. The quantitative contrast involves intact cohorts differing in size, year and composition. Fast-loop evidence rests on the team's own account rather than dated logs. Evidence comes from two modules in one institution, and the authors are insider researchers. Future work should institute dated records of within-semester change, test the framework cross-institutionally with teams holding no design stake, and track graduate outcomes longitudinally (Jackson & Wilton, 2017).

## Conclusion

Embedding live-client briefs into assessed curriculum was associated with higher reported relevance and created structured opportunities to enact and evidence applied competencies in

professionally legible forms. Those benefits were conditional, bounded by client-task fit, client capacity, workload proportionality and scaffolding lead time. BRIDGE is offered as an ordered process framework whose principal advance is the dual-tempo specification of evaluation, evidenced asymmetrically and reported as such. The study measures neither employability nor employment, and asks to be judged on the terms it sets.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this article, OpenAI's ChatGPT (GPT-5.3) was used solely to improve the language and readability of the manuscript by helping the authors condense and refine their own descriptions of the research findings. It was not used to generate, analyse or interpret data, or to develop the study's findings and conclusions. All AI-assisted text was reviewed, verified and edited by the authors, who take full responsibility for the content of the article.

## References

Aaker, D. A., & Keller, K. L. (1990). Consumer evaluations of brand extensions. *Journal of Marketing, 54*(1), 27–41.

Ankrah, S., & Al-Tabbaa, O. (2015). Universities-industry collaboration: A systematic review. *Scandinavian Journal of Management, 31*(3), 387–408.

Argyris, C., & Schon, D. A. (1978). *Organizational learning: A theory of action perspective. Addison-Wesley.*

Ashford-Rowe, K., Herrington, J., & Brown, C. (2014). Establishing the critical elements that determine authentic assessment. *Assessment & Evaluation in Higher Education, 39*(2), 205–222.

Bandaranaike, S., & Willison, J. (2015). Building capacity for work-readiness: Bridging the cognitive and affective domains. *Asia-Pacific Journal of Cooperative Education, 16*(3), 223–233.

Bennett, D. (2019). Graduate employability and higher education: Past, present and future. *HERDSA Review of Higher Education, 5*, 31–61.

Bennett, D., Knight, E., Divan, A., & Bell, K. (2019). Marketing graduate employability: The language of employability in higher education. In J. Higgs, G. Crisp, & W. Letts (Eds.), *Education for employability* (Vol. 2, pp. 105–116). Brill Sense.

Biggs, J. B., & Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). Open University Press.

Billett, S. (2009). Realising the educational worth of integrating work experiences in higher education. *Studies in Higher Education, 34*(7), 827–843.

Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education, 38*(6), 698–712.

Bovill, C. (2020). *Co-creating learning and teaching: Towards relational pedagogy in higher education*. Critical Publishing.

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*(4), 589–597.

Bridgstock, R. (2009). The graduate attributes we've overlooked: Enhancing graduate employability through career management skills. *Higher Education Research & Development, 28*(1), 31–44.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325.

Clarke, M. (2018). Rethinking graduate employability: The role of capital, individual attributes and context. *Studies in Higher Education, 43*(11), 1923–1937.

Coates, H., Zlatkin-Troitschanskaia, O., & Pant, H. A. (2016). Assessing student learning outcomes in higher education: Challenges and international perspectives. *Assessment & Evaluation in Higher Education, 41*(5), 655–661.

Cooper, L., Orrell, J., & Bowden, M. (2010). *Work integrated learning: A guide to effective practice*. Routledge.

Corley, K. G., & Gioia, D. A. (2011). Building theory about theory building: What constitutes a theoretical contribution? *Academy of Management Review, 36*(1), 12–32.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage.

Eimer, A., & Bohndick, C. (2023). Employability models for higher education: A systematic literature review and analysis. *Social Sciences & Humanities Open, 8*(1), Article 100588.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review, 14*(4), 532–550.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal, 50*(1), 25–32.

Ferns, S., Russell, L., & Kay, J. (2016). Enhancing industry engagement with work-integrated learning: Capacity building for industry partners. *Asia-Pacific Journal of Cooperative Education, 17*(4), 363–375.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. Sage.

Gibbs, G. (1988). *Learning by doing: A guide to teaching and learning methods*. Further Education Unit, Oxford Polytechnic.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods, 16*(1), 15–31.

Grewal, D., Guha, A., Satornino, C. B., & Becker, M. (2025). The future of marketing and marketing education. *Journal of Marketing Education, 47*(1), 61–77.

Harvey, L. (2001). Defining and measuring employability. *Quality in Higher Education, 7*(2), 97–109.

Healey, M., Flint, A., & Harrington, K. (2014). *Engagement through partnership: Students as partners in learning and teaching in higher education*. Higher Education Academy.

Helyer, R., & Lee, D. (2014). The role of work experience in the future employability of higher education graduates. *Higher Education Quarterly, 68*(3), 348–372.

Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development, 48*(3), 23–48.

Hinchliffe, G. W., & Jolly, A. (2011). Graduate identity and employability. *British Educational Research Journal, 37*(4), 563–584.

Ho, T. T. H., Le, V. H., Nguyen, D. T., Nguyen, C. T. P., & Nguyen, H. T. T. (2023). Effects of career development learning on students' perceived employability: A longitudinal study. *Higher Education, 86*, 297–315.

Holmes, L. (2013). Competing perspectives on graduate employability: Possession, position or process? *Studies in Higher Education, 38*(4), 538–554.

Huang, Y. (2025). How entrepreneurship education influences university graduate employability. *Humanities and Social Sciences Communications, 12*, Article 1108.

Jackson, D. (2015). Employability skill development in work-integrated learning: Barriers and best practice. *Studies in Higher Education, 40*(2), 350–367.

Jackson, D. (2016). Re-conceptualising graduate employability: The importance of pre-professional identity. *Higher Education Research & Development, 35*(5), 925–939.

Jackson, D., Riebe, L., Meek, S., Ogilvie, M., Kuilboer, A., Murphy, L., Collins, N., Lynch, K., & Brock, M. (2023). Using an industry-aligned capabilities framework to effectively assess student performance in non-accredited work-integrated learning contexts. *Teaching in Higher Education, 28*(4), 802–821.

Jackson, D., & Wilton, N. (2017). Perceived employability among undergraduates and the importance of career self-management, work experience and individual characteristics. *Higher Education Research & Development, 36*(4), 747–762.

Janssens, L., Bakkes, S., & De Vries, T. (2022). Do European quality assurance frameworks support transformative learning for sustainable development? *International Journal of Sustainability in Higher Education, 23*(8), 1632–1652.

Kalfa, S., & Taksa, L. (2015). Cultural capital in business higher education: Reconsidering the graduate attributes movement and the focus on employability. *Studies in Higher Education, 40*(4), 580–595.

Kemper, J. A., & Ballantine, P. W. (2019). What do we mean by sustainability marketing? *Journal of Marketing Management, 35*(3-4), 277–309.

Kinash, S., Crane, L., Judd, M. M., & Knight, C. (2016). Discrepant stakeholder perspectives on graduate employability strategies. *Higher Education Research & Development, 35*(5), 951–967.

Knight, P., & Yorke, M. (2003). *Assessment, learning and employability*. Society for Research into Higher Education & Open University Press.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.

Laasch, O., Moosmayer, D. C., & Antonacopoulou, E. P. (2023). The interdisciplinary responsible management competence framework: An integrative review of ethics, responsibility, and sustainability competences. *Journal of Business Ethics, 187*(4), 733–757.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Leal Filho, W., Shiel, C., Paco, A., Mifsud, M., Avila, L. V., Brandli, L. L., Molthan-Hill, P., Pace, P., Azeiteiro, U. M., Vargas, V. R., Caeiro, S., & Delgado, C. (2019). Sustainable development goals and sustainability teaching at universities: Falling behind or getting ahead of the pack? *Journal of Cleaner Production, 232*, 285–294.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1–13.

Nowotny, H., Scott, P., & Gibbons, M. (2001). *Re-thinking science: Knowledge and the public in an age of uncertainty*. Polity Press.

Oliver, B. (2015). Redefining graduate employability and work-integrated learning: Proposals for effective higher education in disrupted economies. *Journal of Teaching and Learning for Graduate Employability, 6*(1), 56–65.

Patrick, C.-J., Peach, D., Pocknee, C., Webb, F., Fletcher, M., & Pretto, G. (2008). *The WIL (work integrated learning) report: A national scoping study*. Queensland University of Technology.

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Sage.

Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Brostrom, A., D'Este, P., Fini, R., Geuna, A., Grimaldi, R., Hughes, A., Krabel, S., Kitson, M., Llerena, P., Lissoni, F., Salter, A., & Sobrero, M. (2013). Academic engagement and commercialisation: A review of the literature on university-industry relations. *Research Policy, 42*(2), 423–442.

Richardson, J., & Henschke, K. (2010). Formalising industry feedback in assessment practice for work-integrated learning. *The International Journal of Knowledge, Culture and Change Management, 10*(5), 73–84.

Rowe, A. D., & Zegwaard, K. E. (2017). Developing graduate employability skills and attributes: Curriculum enhancement through work-integrated learning. *Asia-Pacific Journal of Cooperative Education, 18*(2), 87–99.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144.

Schon, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Shanahan, D. E., Palmer, L. H., & Salas, J. (2021). Achieving scaled and sustained use of client-based projects in business school marketing education: A proposed suprastructure. *Journal of Marketing Education, 43*(1), 59–74.

Smith, C. (2012). Evaluating the quality of work-integrated learning curricula: A comprehensive framework. *Higher Education Research & Development, 31*(2), 247–262.

Stake, R. E. (1995). *The art of case study research*. Sage.

Succi, C., & Canovi, M. (2020). Soft skills to enhance graduate employability: Comparing students' and employers' perceptions. *Studies in Higher Education, 45*(9), 1834–1847.

Tight, M. (2023). Employability: A core role of higher education? *Research in Post-Compulsory Education, 28*(4), 551–571.

Tomlinson, M. (2017). Forms of graduate capital and their relationship to graduate employability. *Education + Training, 59*(4), 338–352.

Tracy, S. J. (2010). Qualitative quality: Eight “big-tent” criteria for excellent qualitative research. *Qualitative Inquiry, 16*(10), 837–851.

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management, 14*(3), 207–222.

Trede, F. (2012). Role of work-integrated learning in developing professionalism and professional identity. *Asia-Pacific Journal of Cooperative Education, 13*(3), 159–167.

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education, 43*(5), 840–854.

Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. Cambridge University Press.

Whittaker, J. G. (2016). Skills gap: A strategy for increasing knowledge worker supply and demand. *Journal of Business, 1*(4), 13–24.

Wu, Y., Ross, A., & Stewart, M. (2025). AI-enabled industry-collaborative assessment (AI-ICAM): Advancing pedagogy and university-industry collaboration. In *The European Conference on Education 2025: Official conference proceedings* (pp. 467–487). IAFOR.

Wu, Y., & Stewart, M. (2024). Gamified assessment in higher education: A conceptual framework for students' motivation and engagement. In *The European Conference on Education 2024: Official conference proceedings* (pp. 553–566). IAFOR.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage.

Yorke, M., & Knight, P. T. (2006). *Embedding employability into the curriculum*. Higher Education Academy.

Zainudden, D., Broom, M., Nousek-McGregor, A., Stubbs, F., & Veitch, N. (2022). Embedding 21st-century employability into assessment and feedback practice through a student-staff partnership. *Access Microbiology, 4*(3), Article 000329.

**Contact email:** ywu002@dundee.ac.uk

# Designing AI-Supported Assessment in Marketing Education: Specified Latitude and the CRAFT Framework

Ying Wu, University of Dundee, United Kingdom
Andrew Ross, University of Dundee, United Kingdom
Malcolm Stewart, University of Stirling, United Kingdom
Fahima Alam, Queen Mary University of London, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Universities must prepare marketing graduates for practice in which generative artificial intelligence (AI) is routine, while preserving assessment validity and independent disciplinary judgement. Detection-led prohibition is difficult to enforce reliably, and permission without evidential requirements can weaken assessability, yet the literature does not specify the design conditions under which authentic assessment converts permitted AI participation into demonstrable evaluative judgement. This study examines two iterations of an undergraduate marketing module, with a postgraduate module as an embedded contrast, across 191 enrolled and 182 marked students. It combines a documentary audit of review forms, briefs and rubrics with a documentary census of 182 submissions across four assessment sets, analysed abductively against stated criteria. Three findings are advanced. Client authenticity supplied constraints against which generated output could be judged, but the assessment architecture specified no AI participation and carried no AI-related criterion in either iteration. Although generative AI was taught in class and demonstrated for specified creative operations, documented declaration rates were 5.8% and 28.6% across the two undergraduate iterations, and of 182 submissions none evidenced revision or disciplinary justification of generated output. Mean marks in the focal undergraduate module did not rise, but no AI-related criterion existed from which an effect of specification could be evaluated. The case is diagnostic rather than evaluative. The study proposes specified latitude, the task-stage-specific mapping of permitted AI operations to the evidence of judgement they must leave, operationalised through CRAFT as a design response to the documented gap.

*Keywords:* generative artificial intelligence, authentic assessment, marketing education, evaluative judgement, assessment design

iafor
The International Academic Forum
www.iafor.org

## Introduction

Generative artificial intelligence (AI) now performs the operations that marketing assessment has long used as evidence of competence. Systems in general use draft copy, generate campaign concepts, produce brand-styled imagery and propose elements of strategic positioning (Cillo & Rubera, 2025; Grewal et al., 2025b; Guha et al., 2024). The consequence is not primarily a policing problem. It is that the artefacts on which marketing assessment relies, the campaign plan, the concept board, the pitch deck, have lost their former standing as unambiguous evidence of the capabilities they were designed to certify. An assessment regime that continues to infer capability from artefact quality alone is making an inference the technology no longer supports, and the defensibility of such inferences is what assessment validity consists in (Ali & Maroulis, 2026; Kane, 2013).

Institutional responses have converged on three positions, none of which resolves the problem. Detection-led prohibition is difficult to enforce reliably and is discouraged by sector guidance (Miao & Holmes, 2023; Quality Assurance Agency for Higher Education [QAA], 2023; Xia et al., 2024). Permission without evidential requirements preserves professional realism but can weaken assessability, leaving no account of what the student contributed. A third position, appeal to authentic assessment, has become the default recommendation, on the reasoning that tasks situated in real professional contexts resist automated completion. Kofinas et al. (2025) contradict it directly: authentic assessments did not safeguard academic integrity, and markers could not reliably detect AI contribution.

Marketing education is theoretically revealing rather than merely convenient. AI already performs central professional operations in the discipline, campaign artefacts are highly generative and multimodal, and artefact quality is therefore unusually weak evidence of unaided capability (Grewal et al., 2025a; Narang et al., 2025). Client constraints, meanwhile, supply natural standards against which generated material can be judged.

What remains unspecified is the design problem itself. Neither banning AI, nor permitting it without condition, nor appending a generic declaration that AI was used somewhere, tells an assessment designer which AI operations to permit, at which point in the task, and what evidential trace that permission must leave. Recent work establishes that evaluative judgement of AI outputs is the capability now at stake (Bearman et al., 2024) and that the digital must be designed into authentic assessment rather than added to it (Nieminen et al., 2023), but neither specifies the architecture that would produce such judgement or render it markable. This study addresses that specification gap.

The aim is to examine how generative AI can be embedded within authentic marketing assessment so that its use becomes visible, evaluable and educationally productive, while independent disciplinary judgement and assessment validity are preserved. Three questions follow. First, what evidential gaps arise when generative AI is pedagogically introduced and demonstrated but not specified within the assessment architecture? Second, what can cross-iteration and embedded-contrast evidence establish, and not establish, about assessment outcomes under an AI-silent architecture? Third, what design conditions follow for making AI-mediated evaluative judgement visible, attributable and assessable, and what bounds them?

The study examines an undergraduate marketing module delivered across two academic years in a Scottish business school, with a postgraduate module providing an embedded contrast. It draws on the formal assessment architecture, a documentary census of 182 submissions across

four assessment sets, reflective teaching records and institutional assessment records. The case is diagnostic rather than evaluative: it documents an architectural absence rather than testing an implemented intervention. Its contribution is to separate disciplinary authenticity from evidential assessability, to separate pedagogical preparation from assessment architecture, and to propose specified latitude and the CRAFT framework as a design response to the gap the case documents.

## Literature Review and Theoretical Development

### Generative AI and the Limits of Prohibition

The early scholarly response to generative AI in higher education treated it as a threat to assessment security, and the resulting literature is organised around detection, deterrence and redesign. Xia et al. (2024), reviewing 32 empirical studies, found detection approaches unable to sustain the evidential burden placed on them and recommended that institutions rethink assessment policy rather than invest further in detection. Sector guidance reaches the same conclusion (Miao & Holmes, 2023; QAA, 2023). The open question is therefore not whether prohibition works, which is largely settled, but what a redesign-led response must contain in order to work.

On that question the literature is thinner and, in one respect, misleading. The redesign most often recommended is a move toward authentic, contextualised tasks, on the assumption that situational specificity defeats general-purpose generation. Kofinas et al. (2025) found it wanting, separating two properties the literature conflates. Authenticity may raise the standard generic output must meet, a quality effect, without making provenance visible, an evidential effect.

A second strand specifies permitted use. Perkins et al. (2024) propose a five-level scale from no AI use to full AI use, attaching expectations to permission: at its collaborative levels students evaluate and modify generated output and supply reflective commentary on how it was questioned and edited. The scale is therefore not merely permissive. Its granularity, however, is the task as a whole. It assigns a level to an assessment rather than mapping permitted operations to stages of an authentic task, and leaves open how the resulting process evidence should enter marking criteria and how the judgement it evidences should be made individually attributable. Pilot evidence shows the scale operating at that task level (Furze et al., 2024). The narrower specification problem is the one addressed here.

### Authentic Assessment in Marketing Education

Authentic assessment asks students to perform tasks resembling professional practice, judged against criteria, audiences and constraints drawn from the field (Herrington & Oliver, 2000; Villarroel et al., 2018). Ashford-Rowe et al. (2014) specify its critical elements as an ill-structured problem, a complete product and an audience beyond the assessor; Boud and Soler (2016) add that the capability should outlast the course. In marketing this is enacted principally through client-based projects, whose effects on applied skill development are reasonably well evidenced (Cummins & Johnson, 2023; Shanahan et al., 2021), a pattern Vlachopoulos and Makri (2024) confirm across disciplines while noting that implementation quality drives outcomes. Classroom studies report students partnering with industry experts on AI-driven campaign work (Richter et al., 2025) and comparing their output against generated alternatives

(Al-Fattal, 2025). These describe pedagogical designs rather than the criteria against which AI-mediated work is marked.

Generative AI unsettles this literature. Authentic assessment derives much of its warrant from the claim that a real context resists retrieval and simulation; if a generative system can produce a plausible campaign concept for a named client after a short prompt, that warrant weakens. Nieminen et al. (2023) argue that the digital dimension of authentic assessment is routinely under-designed, treated as a delivery medium rather than a constitutive feature of the practice simulated. This does not, however, tell the designer what an AI-inclusive task should require.

## Evaluative Judgement Under Generative AI

The most productive theoretical response reframes the capability at stake. Bearman and Ajjawi (2023) argue that because an AI system's reasoning cannot be traced, pedagogy must shift toward judging outputs and processes rather than reconstructing method. Bearman et al. (2024) extend this into evaluative judgement, the capacity to judge the quality of work of self and others (Tai et al., 2018), explaining why a disclosure statement is inadequate: it records that AI was used without evidencing any judgement about what it produced.

Two limitations bound this contribution. It is conceptual, specifying the target capability without the task architecture that would elicit and evidence it. And evaluative judgement of AI output is disciplinary rather than general: judging a generated concept off-brand requires knowledge of the brand, audience and category.

Recent practice-grounded work identifies voice, judgement and taste as human capabilities that remain consequential in AI-mediated marketing work, judgement covering decisions about when generated outputs should be trusted, adapted or rejected (Gonsalves et al., 2026). That advances understanding of what should remain attributable to the learner. It leaves open the question addressed here: how an authentic task can require those judgements to become visible, attributable and markable.

## Research Gap and Theoretical Architecture

Table 1 synthesises these strands against the use made of each here. Together they identify a gap that is architectural rather than empirical. Prohibition is unreliable; authenticity constrains quality without securing provenance; permission scales attach expectations at the level of the task; evaluative judgement names the capability without specifying the design that evidences it. What is missing is an account of how permitted AI participation becomes evidence of student judgement that an assessor can attribute and mark.

The theoretical architecture adopted here is deliberately narrow and ordered. Assessment-validity theory establishes why the artefact-to-capability inference matters and what its defensibility requires (Ali & Maroulis, 2026; Kane, 2013). Authentic assessment explains what supplies the disciplinary and professional constraints against which generated output can be judged. Evaluative judgement identifies the human capability that must remain attributable. Specified latitude, developed below, is proposed as the design property connecting them, and reflective practice supplies supporting vocabulary rather than a fourth pillar.

**Table 1**
*Critical Synthesis of Scholarship Informing the Study*

| Strand | Key sources | Core claim | Limitation for the present problem | Use in this study |
|---|---|---|---|---|
| Assessment validity and inference | Kane (2013); Ali and Maroulis (2026) | Validity rests on the defensibility of inferences drawn from assessment evidence; generative AI destabilises the artefact-to-capability inference | Establishes the problem without specifying the design response | Supplies the central framing of the paper |
| Prohibition and detection | Xia et al. (2024); Miao and Holmes (2023); QAA (2023) | Detection is unreliable; redesign is recommended over prohibition | Establishes what fails without specifying what a redesign must contain | Motivates the design problem |
| Authenticity as an integrity safeguard | Kofinas et al. (2025) | Authentic assessments neither prevented nor revealed AI contribution | Diagnoses failure without separating quality from provenance effects | Motivates the finding on authenticity and specification |
| Authentic assessment design | Herrington and Oliver (2000); Ashford-Rowe et al. (2014); Villarroel et al. (2018); Boud and Soler (2016); Vlachopoulos and Makri (2024) | Specifies the elements of authenticity and of sustainable capability | Pre-dates generative AI; assumes real context resists simulation | Supplies the client-anchored task component |
| Client-based projects in marketing | Shanahan et al. (2021); Cummins and Johnson (2023) | Evidences applied skill development; scaling depends on design not enthusiasm | Silent on AI participation within client tasks | Establishes the disciplinary anchor |
| Permission specification | Perkins et al. (2024); Furze et al. (2024) | Attaches transparency, evaluation and modification expectations to levels of permitted AI use, with pilot implementation evidence | Operates through task-wide permission categories rather than mapping permitted operations to task stages, required evidence, marking criteria or individual attribution | Establishes the narrower specification problem this study addresses |
| Evaluative judgement and irreducibly human capability | Bearman and Ajjawi (2023); Bearman et al. (2024); Tai et al. (2018); Gonsalves et al. (2026) | Judgement of AI outputs is the capability at stake; practitioner evidence identifies voice, judgement and taste as non-delegable | Identifies what should remain attributable without specifying the architecture that evidences it | Supplies the criteria component and the outcome construct |
| AI and creative performance | Urban et al. (2024) | AI assistance improved creative output; metacognitive monitoring did not improve | Experimental and decontextualised | Supports the claim that generation improves more readily than judgement |
| Digital design in authentic assessment | Nieminen et al. (2023) | The digital is routinely under-designed within authentic assessment | Does not specify what an AI-inclusive task should require | Frames the professional-fidelity argument in the discussion |
| Marketing practice and curricula under AI | Guha et al. (2024); Grewal et al. (2025a, 2025b); Narang et al. (2025); Cillo and Rubera (2025) | AI is embedded in marketing practice and entering marketing curricula | Largely agenda-setting rather than design-specifying | Establishes disciplinary urgency and specificity |

*Note.* Developed by the authors.

# Methodology

## Research Design and Case

The study is a reflective, documentary-led case study of a bounded instructional setting (Yin, 2018), examined longitudinally. The change between iterations is treated as a cycle of design, observation and adjustment (Schön, 1983) rather than two unrelated snapshots, and its primary evidence consists of artefacts of design rather than informant accounts. It is not design-based research: no intervention was theorised in advance and no testing protocol pre-specified.

The case is BU31013, an undergraduate module in sustainability and digital marketing, delivered in 2024/25 and 2025/26 (115 enrolled; 104 and 108 marked on the two components). BU52039, a taught postgraduate module in brand management and strategy delivered across the same years (76 enrolled; 74 marked on the report component), provides an embedded contrast in which live-client presence varies while its assessment format, weighting and word counts are held constant. In the focal undergraduate module the permitted campaign plan word count rose from 1,500 to 2,000 words between iterations. The full study population is 191 enrolled and 182 marked, and all four assessment sets were censused, giving 182 submissions, one for each marked student. One BU31013 2024/25 record was a duplicate file submitted by a student whose work is already represented in the census, and it was removed so that the census carries a single record per marked student. The two focal iterations hold module, level, teaching team, assessment format and client presence constant while the assessment architecture remained silent on AI in both. Table 2 sets out the case, its design and the evidence available.

## Data Sources

Four sources are used. Module assessment review forms, briefs and marking criteria for both iterations document what was formally required and rewarded, and are authoritative for the assessment architecture. All 182 submissions across the four assessment sets were audited for documentary evidence of AI participation, using predefined search terms covering generative tools and disclosure phrasing; 178 were text-searchable, and four non-extractable files are treated conservatively as carrying no declaration. Every match was then read in context, and matches evidencing only a client brand’s own AI features, or arising from reference-list URL artefacts, were excluded. AI use was never inferred from writing style, image quality or detection software. Reflective teaching records supply the rationale for design decisions and are treated as artefacts authored by parties with a stake in the account. Institutional assessment records supply mark distributions by component and iteration on the University's 23-point scale.

Two limits on this evidence must be stated. Student feedback was available as contextual evidence concerning the module and assessment experience, but did not specifically address generative AI; no claim about students' subjective experience of AI is therefore advanced. Second, the submission audit establishes what students documented, not what they did, and is best read as a measure of the architecture's evidential yield rather than of behaviour.

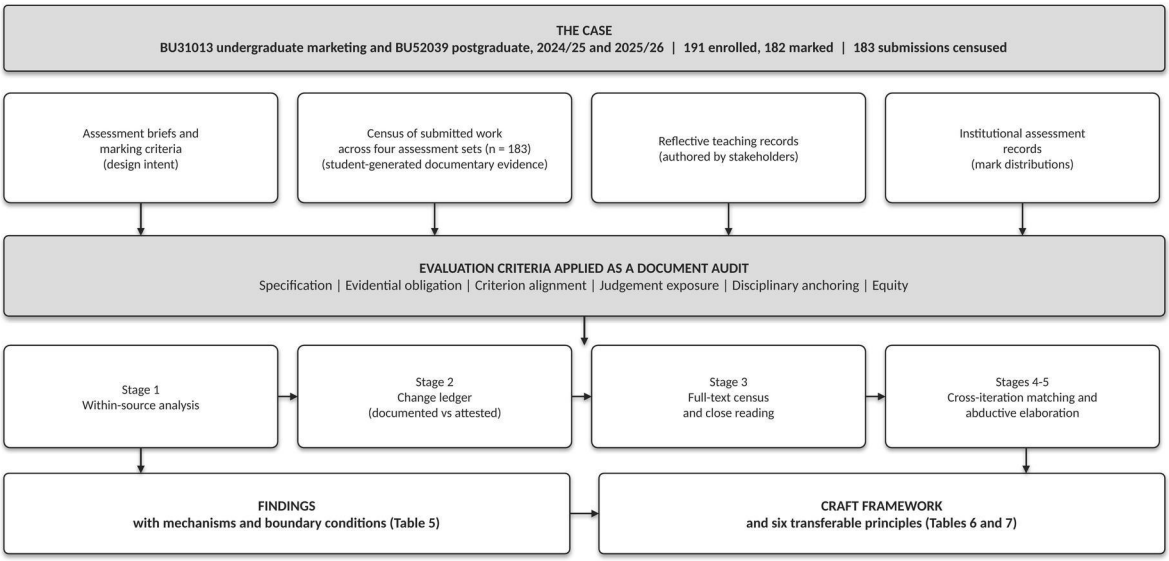
**Table 2**

*Case Context, Assessment Architecture and Documented Change Across Iterations*

| Component | 2024/25 | 2025/26 | Available evidence | Analytical relevance |
|---|---|---|---|---|
| Focal module | BU31013, undergraduate, sustainability and digital marketing | Unchanged | Module assessment review forms | Level, content and teaching team held constant |
| Live client | Intelligent Growth Solutions, a business-to-business technology company | Campy Growers, a consumer-facing sustainable growing initiative | Review forms; briefs | Client presence constant; client type varies |
| Assessment structure | Podcast with show notes (30%); SOSTAC campaign plan (70%) | Unchanged | Briefs; review forms | Format and weighting held constant |
| Permitted word count, campaign plan | 1,500 words | 2,000 words (plus or minus 10%) | Review forms | A documented change in task size; a confound for cross-iteration mark comparison |
| AI in the assessment architecture | No reference to generative AI in the review form, brief or marking criteria | Unchanged. No reference to generative AI in the review form, brief or marking criteria | Review forms signed 16 July 2024 and 7 August 2025; briefs; rubrics | The design variable of theoretical interest is absent in both iterations |
| Marking criteria | Five criteria: client strategy research, analysis of marketing effectiveness, proposed campaign, use of evidence, presentation | Unchanged; no criterion concerns AI | Rubrics embedded in the review forms | No credit attaches to documenting or evaluating AI participation |
| AI in teaching | Introduced and discussed in lectures and tutorials | Introduced and demonstrated for podcast music and campaign visual mock-ups | Teaching materials; module team record | Pedagogical preparation delivered without architectural specification |
| Moderation | No moderator assigned; recorded as a new course | Moderated and signed by a named moderator, 7 August 2025 | Review forms | A documented change in assurance process between iterations |
| Embedded contrast module | BU52039, postgraduate: student-selected brand, no live client | BU52039 with a live client; AI appears as a level-of-permitted-use field outside the marking criteria | Review forms; briefs | Varies client presence with format held constant |
| Enrolment | Undergraduate 56; postgraduate 56 | Undergraduate 59; postgraduate 20 | Institutional records | Study population of 191 enrolled, 182 marked |

*Note.* Developed by the authors from module assessment review forms, briefs and marking criteria. Client organisations are named only where their involvement is already a matter of public record.

**Figure 1**
*Research Design and Analytical Process*




*Note.* Developed by the authors. Arrows indicate analytical sequence.

## Data Analysis

Analysis proceeded abductively in five stages (Miles et al., 2014; Patton, 2015), moving between the sensitising concepts developed above and the case material. Stage one analysed each source within itself. Stage two constructed a change ledger recording each documented alteration between iterations alongside what was held constant, classifying each entry by evidential status. Stage three audited all 182 submissions across the four assessment sets against predefined search terms, with every match subsequently examined in context. Stage four applied cross-iteration pattern matching, using the postgraduate contrast as a comparison in which a different variable moved. Stage five elaborated the framework abductively and bounded it through negative-case analysis.

Interpretation was iterative and reflexive rather than treated as a reliability exercise. The lead author conducted the analysis; co-authors acted as critical interpretive partners, challenging theme boundaries and contribution claims. Because the submission audit was a predefined documentary census rather than an interpretive coding exercise, no inter-rater reliability statistic was calculated or claimed.

## Evaluation Criteria

Because assessment design cannot be evaluated without stated criteria, six dimensions were specified in advance of the cross-iteration comparison and applied to each iteration as a document audit rather than a rating of students. Specification asks whether permitted AI operations are named in the brief at the level of the task stage. Evidential obligation asks whether permitted use carries a requirement to produce a trace. Criterion alignment asks whether marking criteria name critical evaluation of AI output with band descriptors. Judgement exposure asks whether the design contains a point at which the student must select, reject or defend without AI mediation. Disciplinary anchoring asks whether the task supplies constraints against which output can be judged wrong. Equity asks whether the design assumes uniform prior AI access or capability, an assumption that survey and interview evidence on perceived risk, trust and tech-savviness gives reason to doubt (Oc et al., 2025); it is treated

throughout as a cross-cutting condition on the other five dimensions rather than as a constitutive component of specified latitude.

### Rigour, Reflexivity and Ethics

Credibility, dependability and confirmability are addressed in the terms set out by Lincoln and Guba (1985). Credibility rests on triangulation across documentary, submission-level and institutional sources, and on reporting the postgraduate contrast, which does not support the study's principal design claim. Dependability rests on the change ledger, on predefined and reportable search terms, and on a stated exclusion rule applied to every match. Confirmability is limited by the authors' position as insider researchers who designed, taught and marked the assessments analysed. Ethical approval was granted by the School of Business Research Ethics Committee, University of Dundee (UoD-SoB-STAFF-2025-16). Submitted work was first audited for the presence or absence of explicit documentary evidence of AI participation; matched passages were then examined in context to characterise the depth of disclosure. No student names, identifiers or marks were linked to submissions, no student-authored text is reproduced, and results are reported only in aggregate.

## Findings

### Authenticity Without Specification

Both iterations of the focal case were anchored in a live external client, and in both the client relationship performed the work the authentic assessment literature predicts: it supplied brand constraints, an audience beyond the marker and a real recipient of the output. The assessment architecture, however, specified nothing about AI. The module assessment review forms signed in July 2024 and August 2025 contain no reference to generative AI, and neither the briefs nor the marking criteria name a permitted AI operation. The 2025/26 campaign plan rubric assesses research into the client's strategy, analysis of marketing effectiveness, the proposed campaign, use of evidence, and presentation. None of its five criteria concerns AI. In the postgraduate module, AI appears only as a level-of-permitted-use field referring students to institutional guidance, positioned outside the marking criteria entirely.

Authenticity and specification operate on different objects. Authenticity acts on output quality, raising the standard generic material must meet before it is credible to a client. Specification acts on provenance, determining whether the origin of that material is recoverable. A task can therefore be highly authentic and wholly unspecified, constraining what students can plausibly submit without establishing what they contributed. This offers one mechanism consistent with Kofinas et al.'s (2025) result: authenticity was not structured to do evidential work. The boundary condition is that the separation is observable only where a task is genuinely ill-structured, since in narrow tasks quality and provenance constraints are harder to distinguish.

### Pedagogical Introduction Without Evidential Obligation

Generative AI was not absent from the module. It was introduced and demonstrated in lectures and tutorials, and students were shown its use for defined creative operations, including music for the podcast and visual mock-ups for the campaign. What did not follow it was any architectural requirement: no brief clause obliged students to submit generated artefacts, document prompts or account for revisions, and no criterion attached credit to doing so. Table 3 reports the result across all four assessment sets. In the demonstrated iteration, 16 of 56

submissions (28.6%) declared AI use and 10 (17.9%) named a tool. In the preceding iteration of the same module, where AI was introduced and discussed but not demonstrated for named operations, the rate was 3 of 52 (5.8%). Descriptively, the declaration rate was roughly five times higher in 2025/26, although this contrast cannot be attributed to the demonstration itself because cohort, client context and the wider AI environment also changed. Seven in ten submissions in that iteration still carried no documentary evidence from which AI participation could be established either way.

The observed pattern is consistent with a mechanism in which the absence of an evidential obligation reduces documentary visibility. A student who did not use AI and one who used it without documenting it are evidentially indistinguishable, leaving the marker no basis for the inference. Teaching students to use generative AI critically is a pedagogical act; making that criticality visible to an assessor is an architectural one, and the first does not deliver the second. Classroom interventions can measurably shift students' confidence in generated output (Beninger et al., 2025), yet such gains remain invisible to assessment unless the architecture obliges a trace. Two boundary conditions apply. Absence of a declaration evidences no documented trace, not absence of use. The keyword method also cannot detect disclosure phrased in unanticipated terms.

**Table 3**

*Census of Documented AI Participation Across Four Assessment Sets*

| Assessment set | Condition at the time of submission | n | Declarations, n (%) | Highest depth of declaration reached |
|---|---|---|---|---|
| BU31013 2024/25 | AI introduced and discussed; architecture silent | 52 | 3 (5.8) | Declaration only |
| BU31013 2025/26 | AI demonstrated for defined operations; architecture silent | 56 | 16 (28.6) | Claimed evaluation, one submission |
| BU52039 2024/25 | AI not taught; no live client; architecture silent | 54 | 0 (0.0) | None |
| BU52039 2025/26 | Level-of-permitted-use field in the brief, outside the marking criteria | 20 | 3 (15.0) | Purpose |
| All four sets | No AI-related marking criterion in any set; one set included a task-level permitted-use field outside the criteria | 182 | 22 (12.1) | No submission reached transformation or disciplinary justification |
| Excluded or non-evidentiary cases | AI as client-brand subject matter (48); reference-list URL artefact only (6); text not extractable (4) | 58 | – | Do not evidence the student's own use |

*Note.* Developed by the authors. Search terms were AI-generated, generative AI, AI-assisted, artificial intelligence, AI use, AI statement, AI declaration, AI was used, used AI and AI tool, together with named tools including ChatGPT, Copilot, Midjourney, Firefly, Gemini, DeepSeek, Perplexity, Claude and Grammarly. Every match was read in context and excluded unless it evidenced the student's own use; the largest excluded category is discussion of a client brand's own AI features, which accounts for 48 submissions and, in the postgraduate 2024/25 set alone, for 35 of 54. AI use was never inferred from writing style, image quality or detection software. Because no brief or criterion required disclosure, all recorded declarations were voluntary. Four submissions could not be text-extracted and are counted conservatively as carrying none.

## Disclosure Depth and the Absence of Attainment Movement

All 22 declarations across the four sets were then read in context and classified by depth. Nineteen stated a purpose and went no further; two acknowledged use without stating one. One evaluated the output, and did so by asserting that a check had been completed rather than reporting what it found. None explained how generated material had been revised or rejected,

and none justified such a decision by reference to the client, the audience, the brand or any marketing criterion. The postgraduate 2025/26 brief is the sharpest test available here, because it carried a level-of-permitted-AI-use field outside the marking criteria. No submission reproduced that field or any statement derived from it. Task-level permission without an attached evidential obligation therefore left no recoverable trace in submitted work. Table 4 reports marks on the University's 23-point scale. Means moved slightly downward between iterations, the podcast from 17.98 to 17.22 (d = -0.31) and the campaign plan from 16.80 to 16.48 (d = -0.09). The postgraduate contrast, where client presence rather than AI treatment varied, shows the larger movement, the individual report rising from 17.18 to 18.85 (d = 0.53).

These figures function as a negative boundary on what can be claimed rather than as a test of the framework. No AI-related criterion existed in either iteration, so the marks cannot evaluate an architecture that was never implemented. Three further changes preclude cross-iteration inference. The client changed from a business-to-business technology company to a consumer-facing organisation, which the companion analysis identifies as affecting how readily students applied taught concepts (Wu et al., 2025). The permitted word count for the campaign plan rose from 1,500 to 2,000 words. Generative tool capability also advanced materially between iterations.

**Table 4**

*Assessment Records by Iteration and Component (23-point scale)*

| Module and iteration | Condition | Component | n | M | SD | Change in M | d |
|---|---|---|---|---|---|---|---|
| BU31013 2024/25 | AI-silent architecture; B2B client; 1,500 words | Podcast (30%) | 50 | 17.98 | 2.08 | - | - |
| BU31013 2025/26 | AI-silent architecture; consumer client; 2,000 words | Podcast (30%) | 54 | 17.22 | 2.77 | -0.76 | -0.31 |
| BU31013 2024/25 | AI-silent architecture; B2B client; 1,500 words | Campaign plan (70%) | 52 | 16.80 | 3.63 | - | - |
| BU31013 2025/26 | AI-silent architecture; consumer client; 2,000 words | Campaign plan (70%) | 56 | 16.48 | 3.65 | -0.32 | -0.09 |
| BU52039 2024/25 | No live client | Group component (30%) | 56 | 19.58 | 2.16 | - | - |
| BU52039 2025/26 | Live client | Group component (30%) | 16 | 20.00 | 1.22 | +0.42 | 0.21 |
| BU52039 2024/25 | No live client | Individual report (70%) | 54 | 17.18 | 3.53 | - | - |
| BU52039 2025/26 | Live client | Individual report (70%) | 20 | 18.85 | 1.67 | +1.67 | 0.53 |

*Note.* M = mean; SD = standard deviation; d = Cohen's d on pooled standard deviations. Marks are on the University of Dundee 23-point scale; percentage conversion is not meaningful on this scale and is not reported. Effect sizes are descriptive, not tests. Cohorts are intact groups differing in size and composition, and no causal inference is drawn. Source: institutional assessment records. Neither undergraduate iteration contained an AI-related criterion; the condition column records what did change.

**Table 5**
*Findings, Evidence, Mechanisms and Boundary Conditions*

| Finding | Empirical basis | Proposed mechanism | Status | Boundary condition |
|---|---|---|---|---|
| Authenticity supplies disciplinary constraints without specifying AI participation | Both iterations client-anchored; neither review form, brief nor rubric refers to generative AI | Authenticity acts on output quality; specification acts on provenance. The two are separable properties of a task | Observed | Separable only where the task is genuinely ill-structured |
| Pedagogical introduction does not itself create an evidential obligation | AI introduced and demonstrated for podcast music and campaign visuals; no brief clause or criterion requires a trace | Teaching criticality is a pedagogical act; making it visible to an assessor is an architectural one | Observed | Teaching record rests partly on the module team's account |
| Voluntary traces yield limited and uneven assessment evidence | Census of 182 submissions across four sets: declaration rates of 5.8%, 28.6%, 0% and 15.0%; 160 (87.9%) carry no declaration | Where evidence is neither required nor rewarded, the assessment yields limited documentary traces | Observed | Absence of a trace is not absence of use; extent of actual AI participation is unknown |
| Declarations evidence use but not evaluative judgement | Of 22 declarations, 19 state a purpose only, 2 acknowledge use only, 1 evaluates the output, 0 explain revision, 0 justify it on disciplinary grounds | A disclosure regime without an evidential obligation produces traces that record use without evidencing judgement about it | Observed; full-text classification of every matched passage in all four sets | Classification rests on what students chose to write; a declaration may understate actual use |
| Attainment data cannot evaluate an architecture that was not implemented | Podcast d = -0.31; campaign plan d = -0.09; postgraduate report d = 0.53 where client presence varied | No AI-related criterion existed in either iteration, so no criterion-linked attainment change was predicted | Negative boundary on claims | Client, permitted word count, tool capability and cohort composition all changed |

*Note.* Developed by the authors. Effect sizes are descriptive, not tests. Status distinguishes what the case observed from what remains to be verified.

# Discussion

## Theoretical Interpretation

The findings address the first question diagnostically: where AI is introduced and demonstrated but not specified, AI-mediated contribution is largely unrecoverable from submitted work. The second is answered as a boundary rather than an effect: outcomes changed little across the undergraduate iterations, but with no AI-related criterion in either year they cannot evaluate specified latitude. The third is addressed prospectively through CRAFT, which states the conditions under which permitted AI participation could be rendered visible, attributable and assessable.

The theoretical significance lies in two separations. First, authenticity supplies disciplinary, client and audience constraints against which generated output can be judged, but those constraints do not tell an assessor which part of the production process belongs to the student. This refines authentic-assessment theory rather than rejecting it, and explains why Kofinas et al. (2025) could find authentic assessments simultaneously demanding and unauditable. Second, pedagogical preparation and evidential architecture are distinct: students were taught

and shown AI use, yet fewer than three in ten left any trace, under an architecture in which nothing required one.

This changes what authenticity means under AI conditions. An authentic marketing task that excludes generative AI is an increasingly incomplete representation of professional practice, extending Nieminen et al.'s (2023) argument from under-designed digital media to professional fidelity. Authenticity and integrity are not in tension as the prohibition literature assumes; they diverge only when authenticity is asked to do evidential work it was never structured to perform. Evaluative judgement also does not follow from AI access: judgement requires something against which output can be judged wrong, and the client's brand, audience and category constraints supply it.

## The CRAFT Framework

The framework proposed here is organised around a single construct. Specified latitude is the task-stage-specific mapping of permitted AI operations to the evidence of the student's own evaluative judgement that those operations must leave behind. So defined it is not a synonym for clearer instructions, nor for a task-level permission category: instructions and permission levels govern what a student may do, whereas specified latitude governs, operation by operation, what the exercise of that permission must make visible and how that visibility enters marking. CRAFT is therefore advanced as a design response to a documented architectural absence, not as a description of an implemented intervention, and its components are stated so that they can be tested and challenged.

CRAFT names five components through which specified latitude is realised. Client-anchored task supplies the disciplinary constraints against which generated output can be judged inadequate; the component is a disciplinary constraint anchor, instantiated here as a live marketing client, with patients, legal fact patterns, engineering specifications, datasets and laboratory conditions as analogous anchors elsewhere. Rule-specified latitude names permitted AI operations by task stage rather than by blanket permission. Attributable trace obliges the exercise of that latitude to leave submitted evidence, such as the generated artefacts alongside the account of their revision. Forced judgement point locates an unmediated moment at which the student selects, rejects and defends. Tuned criteria make critical evaluation of AI output an assessed dimension with band descriptors. The components are interdependent, and the framework's substantive claim concerns their interaction: each fails in a statable way without the others. Table 6 sets out each component with the evidence that would challenge its role.

Three propositions follow, stated so that CRAFT can be falsified rather than merely adopted. First, disciplinary authenticity increases the constraints against which generated output can be judged, but does not itself increase the observability of AI-mediated contribution. Second, permitted AI participation becomes assessable only where task-stage-specific latitude is coupled with an attributable evidential trace and with aligned marking criteria. Third, an individually attributable judgement point strengthens assurance only where it is anchored in disciplinary constraints and connected to documented process evidence. The first is consistent with what this case observed; the second and third are prospective. Equity bounds all three rather than sitting within any one of them: trace requirements, unequal prior AI proficiency and individually performed defence impose unequal burdens, so each component carries an accessibility and support condition (Oc et al., 2025).

## The Exposure Point as a Proposed Assurance Mechanism

The design pairs a written output carrying strategic rationale with a live, student-performed client-facing pitch. Its assurance value is proposed rather than demonstrated: records establish that a pitch occurred, not that individuals defended particular decisions, and group formats complicate attribution. The proposed mechanism is that a defensible account of a creative decision is difficult to hold without having made it, whereas a polished artefact is not.

Locating assurance at such a point rather than in the artefact would align the design with the argument that pedagogy must judge outputs and processes rather than reconstruct method (Bearman & Ajjawi, 2023). Disclosure alone cannot substitute: a disclosure statement is itself an artefact subject to the same generation problem as the work it describes.

Four failure conditions follow. Latitude specified without an attributable trace restores clarity about rules while leaving provenance unrecoverable, the limitation of permission scales used alone. A trace obliged without tuned criteria is submitted but not rewarded, and students economise on it. Criteria tuned without a client anchor leave markers with no disciplinary standard against which evaluation can itself be assessed. A forced judgement point without the other four becomes an integrity check, since there is no documented development to defend.

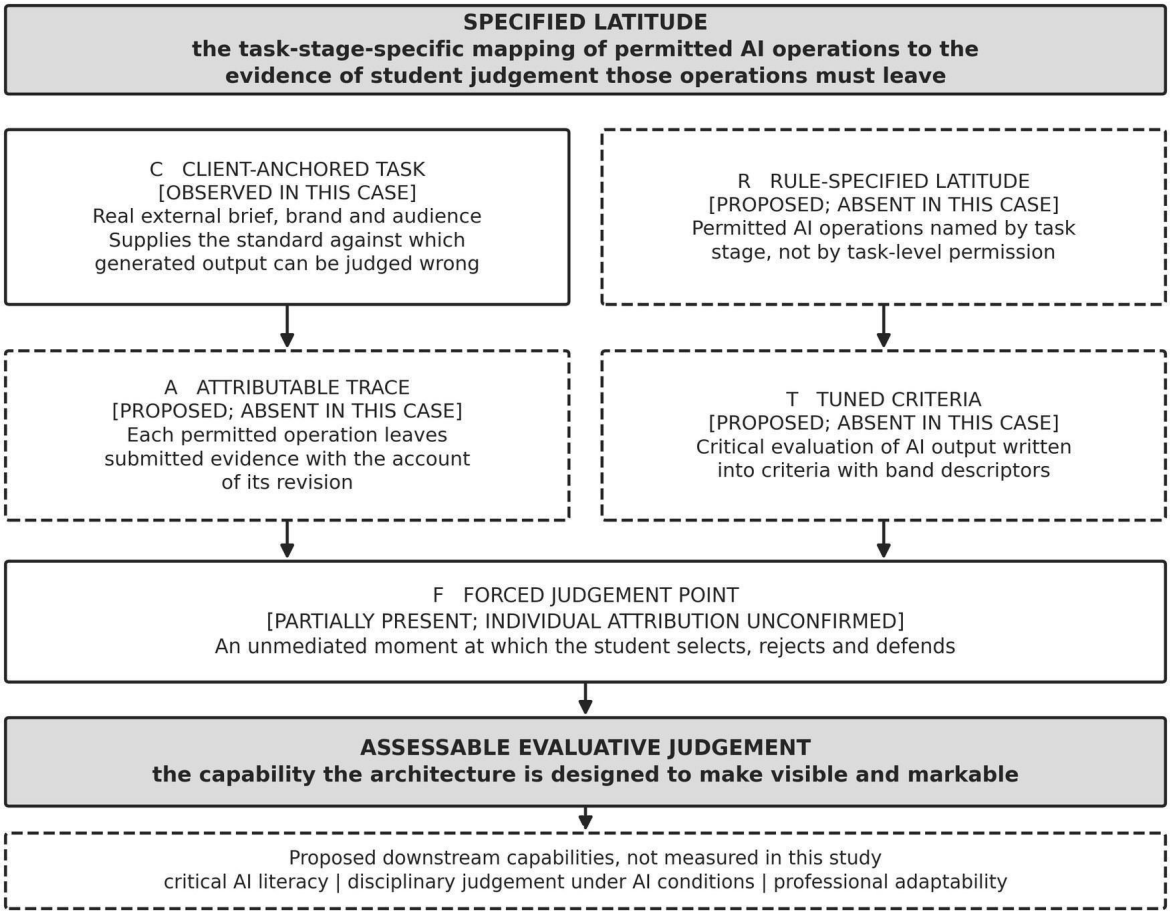
**Table 6**

*CRAFT Components, Status in This Case, and Disconfirming Conditions*

| Component | Design action | Status in this case | Empirical basis or absence | Evidence that would challenge this role |
|---|---|---|---|---|
| C: Client-anchored task | Set the brief with a real external organisation, its constraints and its audience | Observed | Documented in both focal undergraduate iterations; present in BU52039 2025/26 but absent in BU52039 2024/25 | Simulated clients with equivalent constraints producing the same evaluative demand |
| R: Rule-specified latitude | Name permitted AI operations by task stage rather than granting task-level permission | Absent; theoretically proposed | No review form, brief or rubric refers to generative AI in either iteration; where a permitted-use field did exist, in the postgraduate 2025/26 brief, no submission reproduced it | Unspecified tasks yielding equally recoverable accounts of student contribution |
| A: Attributable trace | Oblige each permitted operation to leave submitted evidence alongside the account of its revision | Absent as a requirement; theoretically proposed | No brief clause requires generated artefacts, prompts or revision accounts; 160 of 182 submissions (87.9%) carry no explicit AI-use declaration | Trace requirements fail if assessors cannot distinguish authentic process evidence from AI-generated accounts of that process |
| F: Forced judgement point | Locate an unmediated, individually attributable moment at which the student selects, rejects and defends | Partially present; individual attribution unconfirmed | A live client-facing pitch occurred in both iterations, but records do not establish that individuals defended particular decisions | Written reflection alone producing equivalent assurance |
| T: Tuned criteria | Write critical evaluation of AI output into marking criteria with descriptors at each band | Absent; theoretically proposed | The five campaign plan criteria concern client research, analysis, the proposed campaign, evidence and presentation; none concerns AI | Documented traces attracting equivalent student effort without criterion weighting |

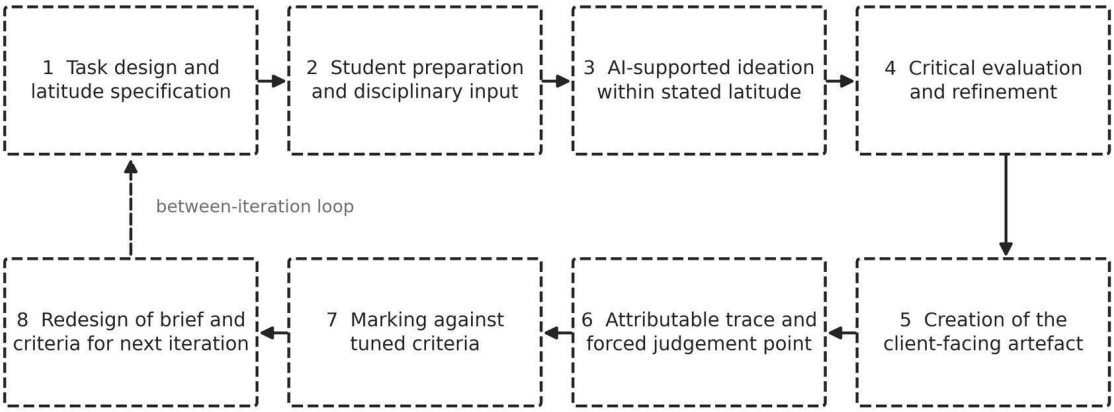
*Note.* Developed by the authors. The case diagnoses the architectural absence; CRAFT proposes the remedy. Only the client anchor was present in the observed iterations, so the framework is advanced for prospective testing rather than as an evaluated intervention.

**Figure 2**
*Specified Latitude and the CRAFT Framework: Status of Each Component in This Case*




*Note.* Developed by the authors. Components are interdependent rather than sequential; failure conditions are specified in Table 6.

**Figure 3**
*Proposed AI-Supported Assessment Redesign Lifecycle*




*Note.* Developed by the authors. Dashed outlines indicate that the lifecycles is the prospective application of CRAFT derived from the diagnostic case. It does not describe an intervention implemented across the two cohorts analysed here.

## Transferable Design Principles

Six principles follow. Table 7 states each with its design action, rationale, intended outcome, the risk it addresses and its boundary condition; the binding constraints are task structure, marker capability, criteria written at band level and staff time.

**Table 7**
*Transferable Principles for Designing AI-Supported Assessment*

| Principle | Design action | Theoretical rationale | Intended outcome | Risk addressed and boundary |
|---|---|---|---|---|
| 1. Begin from a disciplinary problem that supplies constraints | Anchor the task in an external organisation with a brand, an audience and a category | Authentic assessment supplies the standard against which output can be judged inadequate (Villarroel et al., 2018) | Generated output becomes judgeable rather than merely plausible | Addresses generic AI-default output. Fails where the task admits a narrow range of acceptable answers |
| 2. Specify latitude by task stage, not by blanket permission | State which AI operations are permitted at ideation, at production and at refinement | Extends permission scales (Perkins et al., 2024) from level to operation | Ambiguity is removed without prohibition | Addresses undisclosed use arising from unstated scope. Requires briefs to be rewritten, not annotated |
| 3. Attach an evidential obligation to every permission granted | Require generated artefacts to be submitted alongside the account of their revision | Repairs the assessment inference that artefact quality no longer supports | Provenance becomes a submitted object | Addresses unrecoverable authorship. Requires marker capability to judge the account |
| 4. Mark the documented distance, not the polish of the artefact | Write critical evaluation of AI output into criteria with descriptors at each band | Operationalises evaluative judgement of AI outputs (Bearman et al., 2024) | The target capability becomes markable and is signalled to students | Addresses reward for undisclosed generation. Presupposes criteria written at band level |
| 5. Locate an individually attributable exposure point | Pair the documented written output with a live, student-performed defence | Assurance shifts from product originality to accountable judgement (Bearman and Ajjawi, 2023) | Students must hold the reasoning they submit | Addresses outsourced authorship. Staff time cost scales with cohort size; group formats complicate attribution |
| 6. Treat specification as an assessability intervention | Justify redesign by what it makes evidenceable, not by expected marks | Follows from the diagnostic boundary established by the assessment records: attainment cannot evaluate an AI-assessment architecture that is absent from the criteria | Realistic expectations and defensible evaluation of the redesign | Addresses overclaiming. Most likely to be resisted, since it offers no immediate metric improvement |

*Note.* Developed by the authors from the findings reported above. Principles are stated with boundary conditions; transferability depends on those conditions holding in the adopting context.

## Contributions and Implications

Theoretically, the study makes three connected claims. It distinguishes disciplinary authenticity from evidential assessability: authentic client constraints supply standards against which generated output is judged, without making the student's AI-mediated contribution visible. It distinguishes pedagogical preparation from assessment architecture: AI was taught and demonstrated, yet without evidential obligations little documentary visibility followed. And it develops specified latitude, the task-stage-specific connection between permitted AI operations and the evidence of judgement they must leave, operationalised through CRAFT and stated as

three falsifiable propositions. The contribution is one of theory refinement: specified latitude extends authentic-assessment and evaluative-judgement accounts by specifying the evidential conditions under which AI-mediated work can support defensible inferences about student capability.

Empirically, the study contributes a four-set census of 182 submissions in which pedagogical introduction of generative AI coexisted with its absence from the assessment architecture. Its most useful result is negative: across two module codes and two years, no submission evidenced revision or disciplinary justification of generated output, and the iteration carrying an explicit permitted-use field left no trace of it.

Methodologically, the study shows that iterative assessment design can be analysed through a change ledger distinguishing documented from attested change, combined with a documentary census of submitted work against predefined terms. Pedagogically, it supplies six principles and a framework whose disconfirming conditions are stated.

For policy, three implications follow. Task-level permission guidance is insufficient where validity depends on making judgement attributable; guidance should let module teams map permission to discipline-specific evidential requirements. Integrity procedures premised on artefact originality address a property inspection cannot reliably establish, and are better reoriented toward whether a student can account for documented development. Moderation needs criteria naming critical evaluation of AI output, and markers equipped to apply them.

## Limitations and Future Research

Four limitations bound interpretation. The audit establishes documented AI participation, not actual use; declarations were classified from what students wrote, and a submission may understate its own AI use, as one submission containing an apparent unedited generator artefact suggests. Four submissions in a format the extraction pipeline could not open are counted conservatively as carrying no declaration. Second, no AI-specific student feedback exists, so no claim about students' experience of AI is advanced. Third, the authors designed, taught and marked the assessments analysed. Fourth, because no element of the proposed architecture beyond the client anchor was implemented, the case can diagnose the gap but cannot evaluate the remedy; a prospective implementation in which the remaining components are built into the assessment and tested is the study this one indicates.

Four lines of work follow: the student-voice evidence this study lacks, through purpose-designed reflective instruments; multi-institutional implementation testing the three propositions by teams holding no stake in the design; validated measures of evaluative judgement, since attainment on existing criteria is the wrong instrument; and whether specified latitude advantages students with prior AI fluency.

## Conclusion

Generative AI has weakened the inference from artefact to capability on which marketing assessment depended, and the defensibility of that inference is what assessment validity consists in (Ali & Maroulis, 2026; Kane, 2013). This study examined two iterations of a client-anchored undergraduate marketing module, with a postgraduate contrast, and found that client authenticity supplied constraints against which generated output could be judged while the assessment architecture specified no AI participation and carried no AI-related criterion in

either year. Generative AI was nonetheless taught and demonstrated for defined creative operations, yet fewer than three in ten submitted campaign plans contained an explicit AI-use trace. Mean marks in the focal undergraduate module did not rise, but because no AI criterion existed they cannot evaluate an architecture that was never implemented. The paper proposes specified latitude, the task-stage-specific connection between permitted AI operations and the evidence of human judgement those operations must leave, operationalised through CRAFT. The case identifies a potentially transferable design problem requiring testing beyond this setting.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, OpenAI's ChatGPT (GPT-5.3) was used only to support language refinement and improve the clarity, concision and readability of text drafted by the authors, including the presentation of the research findings. The tool was not used to generate, analyse or interpret the study data, or to determine the findings or conclusions. All AI-assisted revisions were subsequently reviewed, verified and edited by the authors, who take full responsibility for the final content of the manuscript.

## References

Al-Fattal, A. (2025). You do it, AI does it, you compare and reflect: Exploring reflective learning with generative AI in principles of marketing. *Journal of Marketing Education*. Advance online publication. https://doi.org/10.1177/02734753251356691

Ali, R., & Maroulis, J. (2026). Assessment validity in the age of generative artificial intelligence: A critical review. *Assessment & Evaluation in Higher Education*. Advance online publication. https://doi.org/10.1080/02602938.2026.2683046

Ashford-Rowe, K., Herrington, J., & Brown, C. (2014). Establishing the critical elements that determine authentic assessment. *Assessment & Evaluation in Higher Education, 39*(2), 205–222. https://doi.org/10.1080/02602938.2013.819566

Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology, 54*(5), 1160–1173. https://doi.org/10.1111/bjet.13337

Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education, 49*(6), 893–905. https://doi.org/10.1080/02602938.2024.2335321

Beninger, S., Reppel, A., Stanton, J., & Watson, F. (2025). Facilitating generative AI literacy in the face of evolving technology: Interventions in marketing classrooms. *Journal of Marketing Education, 47*(2), 112–125. https://doi.org/10.1177/02734753251316569

Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education, 41*(3), 400–413. https://doi.org/10.1080/02602938.2015.1018133

Cillo, P., & Rubera, G. (2025). Generative AI in innovation and marketing processes: A roadmap of research opportunities. *Journal of the Academy of Marketing Science, 53*(3), 684–701. https://doi.org/10.1007/s11747-024-01044-7

Cummins, S., & Johnson, J. S. (2023). The impact of live cases on student skill development in marketing courses. *Journal of Marketing Education, 45*(1), 55–69. https://doi.org/10.1177/02734753211034553

Furze, L., Perkins, M., Roe, J., & MacVaugh, J. (2024). The AI Assessment Scale (AIAS) in action: A pilot implementation of GenAI-supported assessment. *Australasian Journal of Educational Technology, 40*(4), 38–55. https://doi.org/10.14742/ajet.9434

Gonsalves, C., Van Vuren, S., & Gill-Simmen, L. (2026). Assessing the irreducibly human in marketing education: A voice, judgement and taste framework for the generative AI era. *Journal of Marketing Education*. Advance online publication. https://doi.org/10.1177/02734753261454375

Grewal, D., Guha, A., Satornino, C. B., & Becker, M. (2025a). The future of marketing and marketing education. *Journal of Marketing Education, 47*(1), 61–77. https://doi.org/10.1177/02734753241269838

Grewal, D., Satornino, C. B., Davenport, T., & Guha, A. (2025b). How generative AI is shaping the future of marketing. *Journal of the Academy of Marketing Science, 53*, 702–722. https://doi.org/10.1007/s11747-024-01064-3

Guha, A., Grewal, D., & Atlas, S. M. (2024). Generative AI and marketing education: What the future holds. *Journal of Marketing Education, 46*(1), 6–17. https://doi.org/10.1177/02734753231215436

Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development, 48*(3), 23–48. https://doi.org/10.1007/BF02319856

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73. https://doi.org/10.1111/jedm.12000

Kofinas, A. K., Tsay, C. H.-H., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology, 56*(6), 2522–2549. https://doi.org/10.1111/bjet.13585

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.

Narang, U., Sachdev, V., & Liu, R. (2025). When AI wears many hats: The role of generative artificial intelligence in marketing education. *Journal of Public Policy & Marketing, 44*(3), 473–489. https://doi.org/10.1177/07439156251328237

Nieminen, J. H., Bearman, M., & Ajjawi, R. (2023). Designing the digital in authentic assessment: Is it fit for purpose? *Assessment & Evaluation in Higher Education, 48*(4), 529–543. https://doi.org/10.1080/02602938.2022.2089627

Oc, Y., Gonsalves, C., & Quamina, L. T. (2025). Generative AI in higher education assessments: Examining risk and tech-savviness on student's adoption. *Journal of Marketing Education, 47*(2), 138-155. https://doi.org/10.1177/02734753241302459

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Sage.

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The artificial intelligence assessment scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice, 21*(6). https://doi.org/10.53761/q3azde36

Quality Assurance Agency for Higher Education. (2023). *Reconsidering assessment for the ChatGPT era: QAA advice on developing sustainable assessment strategies*. QAA.

Richter, S., Giroux, M., Piven, I., Sima, H., & Dodd, P. (2025). A constructivist approach to integrating AI in marketing education: Bridging theory and practice. *Journal of Marketing Education, 47*(2), 94–111. https://doi.org/10.1177/02734753241288876

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Shanahan, D. E., Palmer, L. H., & Salas, J. (2021). Achieving scaled and sustained use of client-based projects in business school marketing education: A proposed suprastructure. *Journal of Marketing Education, 43*(1), 59–74. https://doi.org/10.1177/0273475319881179

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education, 76*(3), 467–481. https://doi.org/10.1007/s10734-017-0220-3

Urban, M., Dechterenko, F., Lukavsky, J., Hrabalova, V., Svacha, F., Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. *Computers & Education, 215*, Article 105031. https://doi.org/10.1016/j.compedu.2024.105031

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education, 43*(5), 840–854. https://doi.org/10.1080/02602938.2017.1412396

Vlachopoulos, D., & Makri, A. (2024). A systematic literature review on authentic assessment in higher education: Best practices for the development of 21st century skills, and policy considerations. *Studies in Educational Evaluation, 83*, Article 101425. https://doi.org/10.1016/j.stueduc.2024.101425

Wu, Y., Ross, A., & Stewart, M. (2025). AI-enabled industry-collaborative assessment (AI-ICAM): Advancing pedagogy and university-industry collaboration. In *The European Conference on Education 2025: Official conference proceedings* (pp. 467–487). IAFOR. https://doi.org/10.22492/issn.2188-1162.2025.38

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education, 21*, Article 40. https://doi.org/10.1186/s41239-024-00468-z

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage.

**Contact email:** ywu002@dundee.ac.uk

# Beyond Exposure: Emotional Attachment in Heritage Language Retention Among Bilingual Families

Venera Kildyusheva, Campus CA, Türkiye
Anna Volkova, Moscow City University, Russia

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Heritage language retention in bilingual children has traditionally been attributed to factors such as exposure, schooling context, and community practices. Emerging interdisciplinary research suggests that emotional attachment also influences children's language preferences and identity formation (Ghimenton, Cohen, & Minniear, 2023; Vidrine-Isbell, 2017). This study aims to evaluate the role of attachment dynamics in heritage language use among bilingual children from migrant, single-parent households. This research was motivated by the observation of a recurring phenomenon: a bilingual child consistently preferred the heritage language despite full immersion in the dominant societal language. This led to a systematic review of literature on bilingual development, attachment theory, and heritage language maintenance. The review revealed a significant conceptual gap: although family structure is occasionally addressed in heritage language research, single-parent contexts remain largely underexplored. Notably, only one empirical study (Vorobeva, 2021) has examined heritage language transmission in single-parent families without incorporating emotional attachment as an analytical factor. To address this gap, the study presents an exploratory qualitative case analysis of a Russian-Turkish bilingual child raised in Turkey by a single mother. Data were collected through semi-structured interviews and attachment-informed projective methods. Findings suggest that the child's preference for Russian is associated with emotionally significant relationships and repeated person-language associations. Recognizing these dynamics may inform more nuanced approaches to heritage language learning, teaching practices, and educational policy in multilingual contexts.

*Keywords:* heritage language retention, emotional attachment, bilingual children, single-parent migrant families

iafor
The International Academic Forum
www.iafor.org

## Introduction

Language acts not simply as a channel of communication but also as a context where relationships occur, grow, and flourish. Within families, language is used during caregiving, expressions of love, arguments, reassurance, and emotion regulation. For bilingual children, language means more than fluency or frequency of use. A child may be fluent in the language of society while having another language as a more personal and emotionally significant one. Hence, the languages a child uses may indicate not only linguistic ability but also family history.

This approach is particularly relevant to heritage language (HL) maintenance. Existing research highlights the importance of language exposure and parental language practices (Chen et al., 2021), Family Language Policy (Vorobeva, 2021; Wang, 2023), education and external support (Chen & Ragnarsdóttir, 2014; Yousef & Taylor-Leech, 2018), peer influence (Chen et al., 2021), cultural and linguistic identity (Little, 2023; Wang, 2023; Yousef & Taylor-Leech, 2018), and child agency (Little, 2023; Wang, 2023). Although these elements are important, they cannot explain why some bilingual children maintain their HL even when heavily immersed in the society's language.

Recent studies increasingly suggest that family relationships and emotions play a pivotal role in HL maintenance. Tannenbaum and Howie (2002) were among the first to examine language maintenance through the lens of family dynamics and emotions, demonstrating that HL maintenance cannot be discussed without considering a child's familial experiences. Wang (2023) further advanced the understanding of emotions in HL maintenance by illustrating how emotions such as fear, disappointment, joy, and anxiety influence language maintenance and interact with parental language ideologies.

The findings from our systematic review of emotional attachment in HL maintenance align with previous research. Out of 835 articles, 18 studies were analyzed qualitatively. Positive relationships support better HL engagement. Moreover, language maintenance is often linked to feelings of liking and cultural belonging. According to Vidrine-Isbell (2017), language attachment is key. This approach suggests language acquisition is closely connected to social bonding and personal memories. Research in bilingual socialization confirms this. Ghimenton et al. (2025) showed that family background and history help families form emotional attachment to a language. Linguistically natural behavior is influenced not only by competence but also by interpersonal relationships.

However, family structure remains underexplored in HL research. Kildyusheva's review (2026) shows that although some studies address families with an absent parent, single parenthood is usually treated as background information rather than a primary focus. This issue is significant because in immigrant single-parent households, emotional care and responsibility for HL often rest with one individual. Vorobeva's study (2021) of two Russian-speaking single-parent households in Finland is relevant. The author notes that Family Language Policy (FLP) is not a fixed set of rules but varies with mothers' confidence, experiences, beliefs, and negotiations with their children. However, how attachment influences sustained preference for one language over another remains insufficiently studied.

This is the gap the present research aims to fill. The subject, Mark, is a ten-year-old bilingual child who speaks Russian and Turkish and lives in Turkey, where he attends school. He has developed academic Turkish skills and socializes in the Turkish-speaking society but prefers

speaking Russian. Mark is raised primarily by his Russian-speaking mother, while his Russian-speaking grandmother also occupies an important position in his emotional and family environment. The father participates in family life and upbringing, but since the parents have lived separately for a long time, the child rarely sees him. Mark's sustained preference for Russian therefore presents an apparent contrast with the predominantly Turkish-speaking educational and social environment in which he lives.

This case raises the question of why bilingual children may prefer one language over another, even when they possess fluency in both languages. It is important to recognize that the emotional significance attached to each language during a child's upbringing may help explain this phenomenon. In Mark's case, Russian is associated with family life, while Turkish is linked to the school context.

To provide context, a brief overview of attachment theory is warranted. The theory is grounded in Bowlby's concept of the attachment system (Bowlby, 1969) and Ainsworth's empirical research on patterns of attachment (Ainsworth et al., 1978). Although attachment theory does not explicitly address language, it can be extended to linguistic contexts because caregiving relationships are routinely mediated through verbal interaction, allowing language to become associated with emotional experiences, relational security, and human bonding (Vidrine-Isbell, 2017).

The current research views emotional attachment as a complementary analytical angle alongside exposure, FLP, schooling adjustments, self-identity, and social networks rather than an alternative to other models explaining HL maintenance. The study asks: How can attachment dynamics explain the preference for heritage language among bilingual children with no proficiency challenges in the dominant language? It is important to note the study does not suggest a causal link between attachment style and language preference. A single qualitative case cannot prove that a specific attachment style leads to a linguistic outcome. Instead, the study examines whether information about emotional distance, views on caregiver availability, willingness to seek help, relationships with others, and language-use preferences converge enough to interpret emotional attachment as important in HL research.

To investigate this relationship, the study adopted an attachment-informed qualitative case design combining semi-structured interviewing with narrative, sociometric, and projective procedures. Selected relational prompts adapted from the Child Attachment Interview (CAI; Shmueli-Goetz et al., 2008) were used to explore representations of caregiver relationships and emotional availability. Narrative prompts adapted from the Attachment Story Completion Task (ASCT; Bretherton et al., 1990) examined responses to distress, support, and help-seeking. These were complemented by the "Animal Family" drawing task (Wenger, 2003) and the "My Social Circle" procedure (Andrushchenko, 1999), which provided further insight into perceived family positions and the organization of emotional closeness. As the ASCT prompts were incorporated into an adapted dialogue rather than administered in a fully standardised form, the resulting material is interpreted qualitatively rather than diagnostically.

Overall, this approach enabled the case to be examined from two perspectives: as an emotional proximity map and as a linguistic map linking languages to individuals of emotional significance for the child.

The remainder of the paper is organized as follows. First, the case background and methodological procedures are described in greater detail, including the participant's family

and linguistic context and the attachment-informed assessment methods. The findings are then presented across the sociometric, projective, and narrative tasks, with particular attention to patterns of emotional proximity and language use. The discussion interprets these findings in relation to attachment theory, language attachment, and previous research on heritage language maintenance. Finally, the paper considers the implications and limitations of the case study and outlines directions for further research on emotional attachment and HL retention in bilingual and single-parent migrant families.

## Theoretical Background: Attachment, Emotion, and Language

Attachment theory offers a framework for understanding how emotional significance in interpersonal relationships relates to language use. Bowlby (1969) conceptualized attachment as a behavioral system motivating children to remain close to their caregivers during periods of uncertainty, stress, or separation. Ainsworth and colleagues (1978) subsequently demonstrated that children differ in their strategies for maintaining proximity to caregivers and coping with distress. Although attachments may persist throughout development, it is inaccurate to regard them as fixed character traits. Longitudinal research indicates that both stability and change in attachment can occur, depending on family and relational circumstances (Pinquart et al., 2013; Waters et al., 2000).

The principal attachment patterns relevant to the present study can be summarized as follows:

### Secure Attachment

Children perceive their caregivers as reliable sources of assistance in times of need. This sense of stability enables distressed children to remain close to their caregivers or to explore their environment with confidence (Ainsworth et al., 1978).

### Anxious/Ambivalent or Resistant Attachment

Children with this attachment style may continually seek closeness yet experience distress even in the caregiver's presence. This pattern reflects uncertainty regarding the availability of emotional support rather than simply separation (Ainsworth et al., 1978). While manifestations of this attachment may vary in early childhood, in middle childhood it often indicates heightened concern about relationships or caregiver availability. In middle childhood, related representational features may be captured by the preoccupied classification used in the CAI (Shmueli-Goetz et al., 2008).

### Avoidant Attachment

Children with avoidant attachment may refrain from overtly expressing their needs and may appear self-sufficient during distress, particularly when seeking help from caregivers has not yielded the desired response (Ainsworth et al., 1978).

### Disorganized Attachment

Attachment behaviors in this category may be contradictory, disoriented, or lack a coherent strategy (Main & Solomon, 1990).

For the purposes of the present research, emphasis is placed on the anxious-ambivalent or preoccupied attachment pattern, as its primary component is anxiety regarding the availability of the attachment figure rather than early separation itself. Older children exhibiting this pattern tend to display heightened sensitivity to relationships, concern about emotional availability, and difficulty trusting others in stressful situations (Shmueli-Goetz et al., 2008). Consequently, Mark's early separation should be regarded as part of his developmental background rather than as the determinant of a specific attachment style.

Attachment theory itself does not predict language preference. The association between attachment and language arises because caregiving is consistently mediated through language. Processes such as providing reassurance, expressing affection, preparing for sleep, responding to fear, and resolving conflicts occur through specific voices, words, and languages. Research on language acquisition demonstrates that social interaction influences language learning. For example, Kuhl et al. (2003) found that infants who interacted with Mandarin speakers were able to learn phonemes, whereas those exposed only to recorded language did not achieve similar outcomes. Although this does not directly demonstrate the effect of attachment, it suggests that language input is processed differently when accompanied by social interaction compared to passive exposure.

Vidrine-Isbell's (2017) concept of language attachment offers a more direct connection between relational experiences and the meaning attributed to languages. According to this perspective, language becomes associated with people, emotions, memories, and situations through repeated social interactions. Consequently, languages acquire emotional significance through the relationships established with individuals and institutions that use the language.

Empirical support for the relational approach is found in research on heritage language acquisition. Tannenbaum and Howie (2002) linked HL maintenance to family relationships, while Little (2023) highlighted the complexity of this issue from an emotional perspective. Wang (2023) further examined parental emotions in relation to language ideologies and maintenance practices. As demonstrated in the present study, factors such as emotional environment, parent-child relationships, parental language practices, cultural identity, child agency, peer influences, education, and the sociocultural environment are all related to HL acquisition (Kildyusheva, 2026). In summary, positive relationships facilitate greater success in HL usage, whereas anxiety, frustration, or ambivalence can impede the process.

Studies on bilingual socialization also indicate that family ties can influence which language is perceived as more appropriate. Ghimenton et al. (2025) found that the language practices of bilingual families are shaped by life history and emotional attachments rather than by proficiency alone. Therefore, a language associated with a particular family member acquires specific meaning for family members.

Collectively, these approaches establish a theoretical framework for interpreting Mark’s case. The present research does not assert that anxious-ambivalent attachment causes Mark’s preference for Russian. Rather, it examines whether Russian has acquired emotional significance due to its use in Mark’s relationships with caregivers. In contrast, Turkish is associated with academic and social experiences.

## Case Overview

Mark is a 10-year-old boy with a Russian-Turkish bilingual background who has resided permanently in Turkey and attends a Turkish-medium school. He demonstrates fluency and strong academic proficiency in Turkish, although he prefers Russian. His early language development was primarily shaped by his mother, who consistently spoke Russian and maintained close contact with him since childhood, as well as his maternal grandmother, who also spoke Russian. Mark's parents separated when he was two years old. Since then, his Turkish-speaking father has maintained infrequent contact, primarily through online communication.

At five months of age, Mark experienced a two-month separation from his mother due to her work obligations and was cared for by his grandparents during this period. Although he maintained regular video contact with his mother, retrospective family reports indicate no observable signs of distress resulting from the separation. Retrospective evaluation further suggests that Mark adapted well to new environments in both kindergarten and school. His preference for Russian was evident in early childhood and remained consistent after he began attending a Turkish-speaking school.

At the time of the study, Mark exhibited a high level of social adaptation but demonstrated emotional openness and sensitivity to failure. He sought social validation and required extended periods to adjust to life changes, such as the beginning of a new school year. Additionally, he displayed a pronounced need for his mother's presence and relied on her emotional support, particularly during distressing situations and at night.

To examine the hypothesized relationship between the child's language preference and patterns of communication or personal settings, the study employed methods grounded in attachment theory. Data collection included semi-structured interviews incorporating selected items from the Child Attachment Interview (CAI) and the Attachment Story Completion Task (ASCT), as well as the "Animal Family" and "My Social Circle" procedures.

## Results

### "My Social Circle"

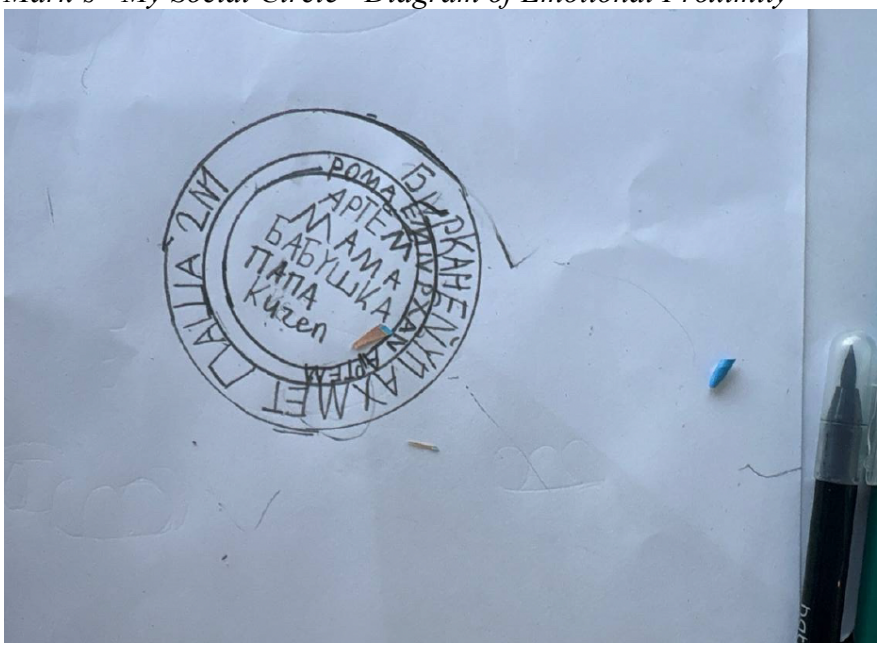
The "My Social Circle" procedure was implemented to assess Mark's perceived emotional proximity to significant individuals within his family and peer environment (Andrushchenko, 1999). Two complementary formats were utilized.

Initially, Mark completed a visual sociometric task that consisted of three concentric circles, each representing a distinct degree of emotional proximity.

- Inner circle: individuals he considered emotionally closest;
- Middle circle: people he liked and regarded as significant but experienced as less immediately close;
- Outer circle: individuals perceived as more distant.

This task was administered privately during an individual session (Figure 1).

**Figure 1**
*Mark's "My Social Circle" Diagram of Emotional Proximity*




Subsequently, Mark was instructed to rank significant individuals in his life on a scale from 5 (greatest perceived closeness) to 1 (least perceived closeness). The list comprised family members, peers, and other significant adults, and the ranking was completed in the presence of his mother. The highest-ranked individuals included his best friend, mother, grandmother, and cousin (Table 1; Figure 2).

**Table 1**
*Integrated Results of the "My Social Circle" Sociometric Diagram and Closeness Rating Task*

| Name | Closeness Circle | Rating (1–5) | Language | Category | Context | Conflict | Additional Notes |
|---|---|---|---|---|---|---|---|
| Mother | Inner | 5 | Russian | Adult | Outside school | No | |
| Father | Inner | 5 | Turkish | Adult | Outside school | No | - |
| Grandmother | Inner | 5 | Russian | Adult | Outside school | No | - |
| Friend of Mother 1 | — | 5 | Russian | Adult | Outside school | Yes | Rated high, not placed in circle |
| Friend of Mother 2 | — | 5 | Russian | Adult | Outside school | Yes | Rated high, not placed in circle |
| Friend of Mother 3 | — | 5 | Russian | Adult | Outside school | Yes | Rated high, not placed in circle |
| Friend of Mother 4 | — | 5 | Russian | Adult | Outside school | Yes | Rated high, not placed in circle |
| Friend of Mother 5 | — | 5 | Russian | Adult | Outside school | Yes | Rated high, not placed in circle |
| German Teacher | — | 5 | Russian | Adult | Outside school | Yes | Rated high, not placed in circle |
| Peer friend Anthony | Inner | 5 | Russian | Peer | School | No | - |
| Cousin | Inner | 5 | Russian | Peer | Outside school | No | - |

| Name | Closeness Circle | Rating (1–5) | Language | Category | Context | Conflict | Additional Notes |
|---|---|---|---|---|---|---|---|
| Peer friend Paul | Outer | 5 | Russian | Peer | School | Yes | Rated high, but placed in outer circle |
| Peer friend Eduard | Middle | 2 | Russian | Peer | School | No | - |
| Peer friend Ronald | Middle | 3 | Russian | Peer | School | No | - |
| Peer friend Alex | Outer | — | Turkish | Peer | School | Yes | Not rated |
| Peer friend Andry | Middle | 4 | Russian | Peer | School | No | - |
| Peer friend Emily | — | 3 | Russian | Peer | Outside school | Yes | Not placed in circle |

A notable discrepancy was observed between the two tasks. Several of his mother's friends received relatively high rankings in the list completed in her presence but were absent from the privately completed sociometric task. Some peers appeared in both tasks, although their relative positions varied. This variation indicates that Mark's representation of social proximity differed depending on the interpersonal context.

A second pattern related to language emerged. Most individuals represented as emotionally close to Mark were Russian-speaking. However, this pattern was not absolute: his Turkish-speaking father was also placed in the inner circle and received the highest closeness rating. Another Turkish-speaking peer was placed in the outer circle and was not included in the ranked table. Emotional proximity therefore did not map exclusively onto Russian, although Russian-speaking relationships predominated across Mark's broader close relational network.

### "Animal Family"

The "Animal Family" drawing task was employed as an additional projective procedure, following Wenger's (2003) methodological description. Mark was invited to represent family members as animals of his choice, without receiving interpretive guidance. Mark depicted himself as a small brown dog and his father as a larger dog, while his mother and grandmother were represented as cats of varying sizes. Other relatives, including cousins, were not assigned animal identities.

Several aspects of the drawing process were noteworthy. Mark repeatedly requested examples of animals, expressing a desire to draw them "correctly," which resulted in the task remaining incomplete. He began with his own figure, drawing it large enough that limited space was left for other family members. He also expressed concern regarding the absence of a white pencil, which prevented him from completing the drawing as intended. Given the projective nature of the procedure, these observations are reported descriptively and are not interpreted as independent indicators of specific psychological characteristics. Their potential significance is considered only in relation to findings from the other tasks.

### ASCT Narratives

The narrative component utilized the ASCT developed by Bretherton et al. (1990) to explore children's representations of attachment-related situations through story completion. As the original procedure was designed for younger children, the prompts were adapted to Mark's age and incorporated into a conversational format rather than administered as a standardized assessment.

The situations presented included a child struggling with homework, a conflict with a friend, and a child afraid to sleep alone in an unfamiliar room. The responses were treated as qualitative narrative material rather than being scored diagnostically.

A recurring pattern emerged across the stories. When protagonists experienced distress, Mark frequently described them as withdrawing or attempting to manage the situation independently, rather than immediately seeking adult support. For example: "He just goes to his room, because… well, he's mad." In another scenario, adult support appeared after an initial period of difficulty: "She [the mother] comes later and helps him with the homework."

Adults, particularly mothers, were generally depicted as eventual sources of support, but their involvement typically followed a period of withdrawal or independent coping. Across the narratives, caregiver assistance was represented as available, though not necessarily immediate.

Taken together, the procedures revealed three recurring patterns: variation in Mark's representation of emotional proximity depending on his mother's presence, a predominantly Russian-speaking network among those represented as emotionally closest to him, and a tendency in the narrative material for distressed protagonists to withdraw or attempt independent coping before caregiver support became available. These patterns are interpreted collectively in the following section rather than as separate diagnostic indicators.

## Interpretation

The findings suggest a relational pattern broadly consistent with anxious-ambivalent or preoccupied attachment tendencies; however, no formal attachment classification is proposed. Across the tasks, Mark exhibited a pronounced inclination toward emotional closeness while expressing uncertainty regarding support arrangements. This pattern may indicate anxious-ambivalent attachment in childhood and a preoccupied attachment representation (Ainsworth et al., 1978; Shmueli-Goetz et al., 2008).

Variations in Mark's responses across the "My Social Circle" activities may reflect his sensitivity to social expectations. He altered his answers when his mother was present, whereas he appeared less influenced by social factors during the privately conducted activity assessing emotional distance. Additionally, analysis of Mark's adapted ASCT narratives revealed that the story's pupils often left or managed difficulties independently before caregivers, particularly mothers, provided assistance. Due to the modified nature of the ASCT, the responses should be interpreted qualitatively.

A notable linguistic pattern emerges in Mark's social environment. Most individuals represented as emotionally close to Mark were Russian-speaking, although his Turkish-speaking father was also positioned in the inner circle and received the highest closeness rating. This suggests that emotional proximity was not exclusively associated with Russian, even though Russian-speaking relationships predominated within his broader network of close relationships. Despite his fluency in Turkish and extensive participation in a Turkish-speaking social environment, Mark's preference for Russian suggests that the language may have acquired particular emotional significance through repeated interactions with his primary Russian-speaking caregivers. Language attachment theory proposes that linguistic experiences become associated with people, emotions, memories, and relational contexts through repeated social interaction (Vidrine-Isbell, 2017). Research on heritage language maintenance likewise

indicates that heritage-language use can be closely connected to family relationships and emotional experience (Little, 2023; Tannenbaum & Howie, 2002).

The concept introduced by Ghimenton et al. (2025) further clarifies this point. Bilingual families often perceive certain languages as natural, a perception shaped by socialization processes, which highlights the influence of emotional attachments. However, the authors did not establish a correlation between attachment type and language use.

This reasoning also applies to Mark's case. While the language of communication may hold emotional significance, routine interactions with family members are shaped by the language used. Consequently, Turkish assumes a more prominent socio-educational role in Mark's life.

## Discussion: Connecting the Case to Broader Research

The case of Mark contributes to the growing body of evidence that heritage-language retention cannot be attributed solely to exposure or linguistic proficiency. The systematic review identified multiple factors influencing heritage language acquisition, including emotional environment, mother-child relationship, identity, parental language practices, individual decisions, peer social networks, and educational context (Kildyusheva, 2026). Mark's consistent use of Russian supports the view that language retention is multidimensional, as Russian remains significant in his key emotional relationships despite his educational and social background being rooted in Turkish.

The relational aspect is particularly significant in this context. Tannenbaum and Howie (2002) associate language retention with family ties, while Little (2023) demonstrates that both family language shift and its reversal are processes with emotional implications that affect personal identity. Findings from the systematic review indicate that maternal affection and positive emotions facilitate heritage language acquisition, whereas fear, anxiety, guilt, and emotional detachment may impede native language retention (Kildyusheva, 2026). It is likely that Mark's primary exposure to Russian occurred within the family, supported by his parents and grandparents.

Vidrine-Isbell (2017) introduces the concept of language attachment, which provides a useful framework for explaining this process. According to this theory, language becomes integrated into networks of individuals and emotions through social communication experiences. Mark's preference for Russian may therefore reflect not only individual language choice but also the relational history accumulated through repeated Russian-language interaction.

Ghimenton et al. (2025) provide a complementary language-socialization perspective. Their study shows that the language experienced as "natural" within bilingual families is shaped by socialization and mobility histories and highlights the role of affective attachment in relationships with particular languages. Their findings do not establish a relationship between attachment style and language preference, but they demonstrate how languages can acquire relational significance through accumulated family experience.

Thus, Mark's preference for Russian exemplifies the connection between individuals and their heritage language. He primarily uses Russian in informal family interactions, whereas his educational background is based in Turkish. The case illustrates that strong affective investment in a heritage language can coexist with extensive institutional exposure to the societal language.

This case contributes to the limited literature on migrant families led by single parents. The systematic review found that single motherhood is rarely analyzed, even when single mothers are present in study samples (Kildyusheva, 2026). Research on single-parent families indicates that single mothers often assume primary responsibility for transmitting language strategies and maintaining emotional continuity within the family. Vorobeva (2021), for example, demonstrates how family language policies in Russian-speaking single-parent families in Finland are shaped by maternal experience, confidence, language ideologies, and everyday mother-child interaction. The current research suggests that when a single parent serves as both the primary emotional figure and the main heritage language speaker, language maintenance acquires an additional dimension.

The findings indicate that attachment theory and research on language socialization are complementary. Attachment theory explains the formation of relationships and the expectations regarding emotional support, while studies on language socialization illustrate how relationships are enacted through language use. Mark's case demonstrates that these processes intersect, as the heritage language acquires a personal significance through repeated use in emotionally meaningful relationships, persisting despite substantial exposure to the dominant societal language.

The results should not be interpreted as definitive evidence that attachment style alone determines language choice. Heritage language maintenance is influenced by multiple factors, including language exposure, education, personal identity, parenting, and social networks. Therefore, Mark's preference for Russian should be viewed not as a contradiction between language environment and language choice, but as an example of the dynamic interaction between these processes.

Several limitations should be considered when interpreting these findings. First, the study is based on a single exploratory case and therefore does not permit generalisation to bilingual children or single-parent migrant families more broadly. Second, the ASCT prompts were adapted and administered conversationally rather than as a standardised assessment, and the projective procedures are interpreted only in conjunction with other qualitative evidence. Third, aspects of Mark's early developmental history rely on retrospective family reports. Finally, the study identifies an interpretive association between relational patterns and language preference rather than a causal relationship. Comparative and longitudinal research is required to determine whether similar emotional-linguistic patterns occur across different family structures, languages, and developmental stages.

## Conclusion

Mark's case demonstrates that the selection of a heritage language encompasses more than the history of language exposure and proficiency. For bilingual children, language is often closely associated with emotional security, continuity of relationships, and a sense of belonging. In Mark's situation, Russian holds particular emotional significance because it is connected to his most important relationships with caregivers. Conversely, Turkish assumes a more prominent role in educational activities and social interactions.

The synthesis of data from sociometric analysis, narratives, and projective tests indicates that Mark's language choices reflect the dynamics of his interpersonal relationships. His predominantly Russian-speaking social circle, perception of caregiver support, and strategies for coping with stress suggest a connection between language use and emotional experience.

While it cannot be conclusively stated that attachment style formation directly influences heritage language choice, it is reasonable to assume that attachment representations and the network of connections between individuals and languages may shape the emotional significance attributed to a heritage language.

By integrating attachment theory with heritage language research, this case study highlights the often-overlooked emotional dimension of heritage language retention. The findings suggest that maintaining a heritage language depends not only on the frequency of its use but also on the manner in which it is experienced, particularly through interpersonal relationships.

The results of this study may inform both future research and educational practice. It is important to provide children with balanced language input while also addressing the emotional and social significance of each language.

## References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the Strange Situation*. Lawrence Erlbaum Associates.

Andrushchenko, T. Y. (1999). Beseda 'Moi krug obshcheniya' v diagnostike sotsial'noi situatsii razvitiya rebenka [The 'My social circle' conversation in diagnosing the social situation of child development]. *Semeinaya psikhologiya i semeinaya terapiya*, (2), 57–64.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment*. Basic Books.

Bretherton, I., Ridgeway, D., & Cassidy, J. (1990). Assessing internal working models of the attachment relationship: An attachment story completion task for 3-year-olds. In M. T. Greenberg, D. Cicchetti, & E. M. Cummings (Eds.), *Attachment in the preschool years: Theory, research, and intervention* (pp. 273–308). University of Chicago Press.

Chen, F., & Ragnarsdóttir, H. (2014). Single-parent immigrant families in Iceland: Lives and educational experiences of their children. *Netla – Online Journal on Pedagogy and Education*. https://doi.org/10.24270/netla.2014/7

Chen, S. H., Zhou, Q., & Uchikoshi, Y. (2021). Heritage language socialization in Chinese American immigrant families: Prospective links to children's heritage language proficiency. *International Journal of Bilingual Education and Bilingualism, 24*(8), 1193–1209. https://doi.org/10.1080/13670050.2018.1547680

Ghimenton, A., Cohen, C., & Minniear, J. N. (2025). Constructing a bilingual French-English habitus through language experience: A socialisation account from children attending a bilingual programme at a state school in France. *Journal of Multilingual and Multicultural Development, 46*(3), 906–922. https://doi.org/10.1080/01434632.2023.2218353

Kildyusheva, V. (2026). A systematic review of emotional attachment and its influence on heritage language retention in single-parent migrant contexts. In T. Guarda, F. Portela, & M. F. Augusto (Eds.), *Advanced research in technologies, information, innovation and sustainability: ARTIIS 2025* (Communications in Computer and Information Science, Vol. 2791, pp. 414–424). Springer. https://doi.org/10.1007/978-3-032-16848-1_31

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100*(15), 9096–9101. https://doi.org/10.1073/pnas.1532872100

Little, S. (2023). 'Half of who you are': Parent and child reflections on the emotional experiences of reversing familial language shift. *International Journal of Bilingualism, 27*(2), 217–231. https://doi.org/10.1177/13670069221125705

Main, M., & Solomon, J. (1990). Procedures for identifying infants as disorganized/disoriented during the Ainsworth Strange Situation. In M. T. Greenberg, D. Cicchetti, & E. M. Cummings (Eds.), *Attachment in the preschool years: Theory, research, and intervention* (pp. 121–160). University of Chicago Press.

Pinquart, M., Feußner, C., & Ahnert, L. (2013). Meta-analytic evidence for stability in attachments from infancy to early adulthood. *Attachment & Human Development, 15*(2), 189–218. https://doi.org/10.1080/14616734.2013.746257

Shmueli-Goetz, Y., Target, M., Fonagy, P., & Datta, A. (2008). The Child Attachment Interview: A psychometric study of reliability and discriminant validity. *Developmental Psychology, 44*(4), 939–956. https://doi.org/10.1037/0012-1649.44.4.939

Tannenbaum, M., & Howie, P. (2002). The association between language maintenance and family relations: Chinese immigrant children in Australia. *Journal of Multilingual and Multicultural Development, 23*(5), 408–424. https://doi.org/10.1080/01434630208666477

Vidrine-Isbell, B. (2017). Language attachment theory: The possibilities of cross-language relationships. In T. M. Zawacki & M. Cox (Eds.), *WAC and second-language writers: Research towards linguistically and culturally inclusive programs and practices* (pp. 95–112). The WAC Clearinghouse. https://doi.org/10.37514/PER-B.2017.0032.2.05

Vorobeva, P. (2021). Hesitant versus confident family language policy: A case of two single-parent families in Finland. *Multilingua, 42*(4), 589–619. https://doi.org/10.1515/multi-2022-0055

Wang, Y. (2023). Speaking Chinese or no breakfast: Emotional challenges and experiences confronting Chinese immigrant families in heritage language maintenance. *International Journal of Bilingualism, 27*(2), 232–250. https://doi.org/10.1177/13670069221126043

Waters, E., Merrick, S., Treboux, D., Crowell, J., & Albersheim, L. (2000). Attachment security in infancy and early adulthood: A twenty-year longitudinal study. *Child Development, 71*(3), 684–689. https://doi.org/10.1111/1467-8624.00176

Wenger, A. L. (2003). *Psikhologicheskie risunochnye testy: Illyustrirovannoe rukovodstvo* [Psychological drawing tests: Illustrated guide]. VLADOS-PRESS.

Yousef, A., & Taylor-Leech, K. (2018). 'It would be nice if someone took the load off you': Arabic-speaking new-Australian mothers and the challenges of heritage language maintenance. *Journal of Home Language Research, 3*, 1–16. https://doi.org/10.16993/jhlr.33

**Contact email:** v.o.kildyusheva@gmail.com

# The Experience of Mexican Female Students in the COIL Mexico-Colombia Project: Insights Into Intercultural Online Collaboration and Educational Exchange

Juana Beatriz Ortiz-Cienega, University of Guanajuato, Mexico
Monica Mondelo-Villasenor, University of Guanajuato, Mexico
Bertha Mondelo-Villasenor, University of Guanajuato, Mexico

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

The COIL Project between Guanajuato (Mexico) and Colombia originated as part of the institutional collaboration between the high schools of the University of Guanajuato and the Colegio Agustiniano Suba in Bogota, Colombia (UG, 2025). This study documents the experience of five female students from the city of Leon, Mexico, distinguished by their have outstanding academic performance and strong collaborative skills. Throughout the development of the Collaborative Online International Learning (COIL) project, the students participated in a virtual master class and engaged in sustained interaction through social networks with peers and teachers from Bogota. These exchanges facilitated the sharing of practices related to environmental care applicable in both countries, as well as the dissemination of cultural traditions and regional celebrations. As a first approach Mexico contributed by presenting the Day of the Dead festivity, while Colombia enriched the collaboration by sharing diverse multicultural activities, including graphic arts, dance, visual arts, and other creative expressions. These exchanges enriched the experience of the Mexican students, who had never traveled abroad and now gained a broader perspective by interacting synchronously with young people of the same age in another country. This COIL project also underscored in the students the recognition that environmental, social, and cultural challenges often exhibit striking similarities across national boundaries. Furthermore, it demonstrated that meaningful international collaboration does not necessarily require substantial financial investment, but rather the commitment of both high schools to join efforts for mutual benefit.

*Keywords:* COIL, Mexico, Colombia, high school, female students

iafor
The International Academic Forum
www.iafor.org

## Introduction

The Escuela de Nivel Medio Superior de Leon (ENMSL) was established in 1878, making it the oldest institution of its kind in the city of Leon, Mexico (Betancourt et al., 2026). At present, ENMSL enrolls 2,986 students, of whom 1,635 are women. A significant proportion of these female students express interest in pursuing careers in science, technology, engineering, and mathematics (STEM). This trend has encouraged ENMSL members to promote additional activities in the fields of science and engineering, as well as to seek opportunities for international collaboration whether in-person or virtual. One notable example is the partnership between the University of Guanajuato and the Colegio Agustiniano Suba in Bogota, Colombia, through a Collaborative Online International Learning (COIL) project (UG, 2025).

This initiative provides ENMSL students, particularly those lacking the financial means to travel abroad, with the opportunity to engage with peers from other countries. It also fosters intercultural exchange and the sharing of diverse perspectives, thereby broadening students' global outlook and enhancing their understanding of contemporary international contexts.

## Literature Review

### Collaborative Online International Learning

The State University of New York Collaborative Online International Learning (SUNY COIL) (2025) describes the Collaborative Online International Learning (COIL) as *an approach that brings students and professors together across cultures to learn, discuss and collaborate as part of their class*. The State University of New York is the pioneer of the COIL virtual exchange model (SUNY COIL, 2025).

According to SUNY COIL (2025), professors design the experience, and students complete the activities designed. COIL can be created in any discipline, encourage active student learning and teamwork with an emphasis on cross-cultural interactions and understanding (SUNY COIL, 2025). A key aspect of COIL is collaboration between professors from two or more higher education institutions from different countries who connect, co-design and co-facilitate collaborative online learning group assignments for students from both institutions (Hackett et al., 2024).

As part of an existing course (SUNY COIL, 2025), the COIL component can take place more commonly as a module that lasts for several weeks during the course. The minimum recommended length for a COIL module is five weeks (SUNY COIL, 2025).

## Methodology

In 2025, when the invitation to work on the COIL project was received, five female students were invited to participate in the ENMSL. They agreed to volunteer, as the project began after the semester was already underway, making it impossible to award a grade for the activity.

The students were between 15 and 16 years old and none of them had previously traveled abroad at that time.

Nevertheless, they expressed interest in visiting countries such as Japan, Sweden, Canada, the United States, Germany, Spain, Colombia, Italy, South Korea, France, or any nation in South America.

All the students intended to pursue careers in Science, Technology, Engineering, and Math fields and come from middle-class backgrounds.

The first COIL activity in which the female students participated consisted of a master class conducted via Google Meet. During this session, the teacher and students gathered in the laboratory of the school to design and construct a water filter using everyday household materials. One of the goals was to raise awareness about the scarcity of potable water in certain regions of both countries and how an affordable tool could be made available to low-income populations. The class was taught by a Colombian professor and was structured around the principles of design thinking. Mexican and Colombian students collaborated synchronously online for 90 minutes.

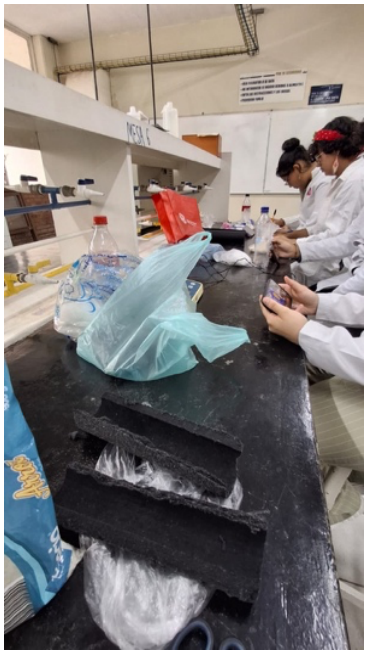
Following the master class, a WhatsApp group was created to enable students to share photographs and information about regional festivities, such as the Day of the Dead in Mexico. This group was kept private to allow the teacher to foster an environment in which the students could express themselves freely.

Additionally, a survey was designed using Microsoft Forms to assess the perspectives and experiences of the five participants. The survey remained open for a period of two days.

## Results

During the master class, the students learn more about collaborative work. Before this activity, they had not had a formal laboratory class.

**Figure 1**
*Working Together From the ENMSL*

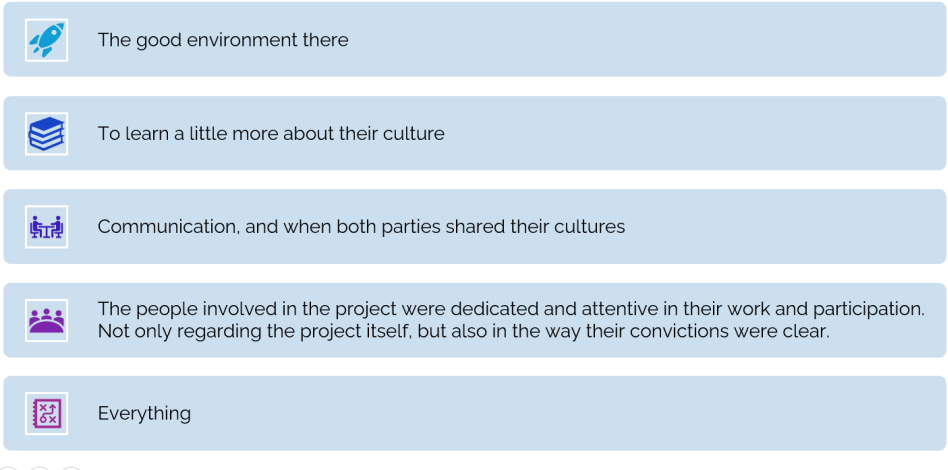


*Note.* Image by author.

Following the review of the survey results provided by Microsoft Forms, valuable data were obtained that offer an initial insight into the impact of the COIL project in the female students from the ENMSL.

As shown in Figure 2, the students said they had a nice time working with Colombian people.

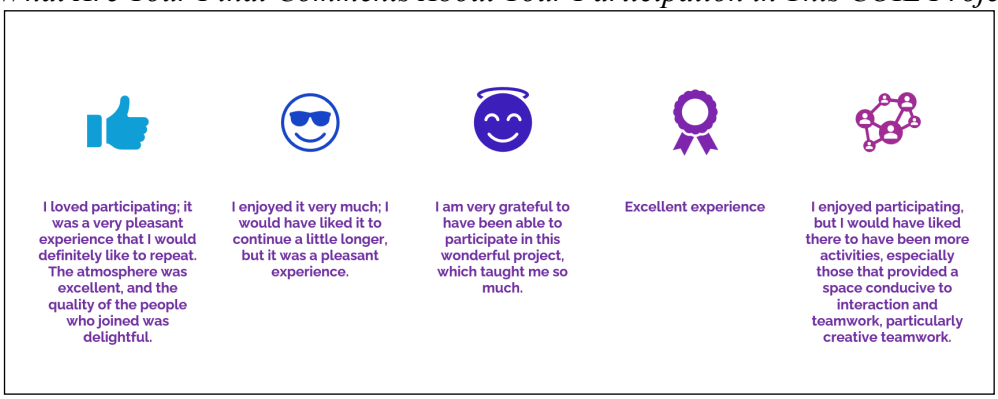
**Figure 2**
*What Did You Like Most About This COIL Project and the Interaction With Colombian People?*




*Note.* Image by author.

The students also commented that they would have loved to continue working on future projects.

**Figure 3**
*What Are Your Final Comments About Your Participation in This COIL Project?*




*Note.* Image by author.

## Conclusion

Unfortunately, the Mexico-Colombia COIL Project is no longer taking place due to the resignation of the head teacher at Colegio Agustiniano Suba. However, the University of Guanajuato is still in contact with the rector of the Colombian institute, awaiting the resumption of the agreement and activities as soon as possible.

For the teacher from the Escuela de Nivel Medio de Leon (ENMSL), this learning experience has been immense: ranging from learning and implementing a new methodology to

witnessing the growth of students who worked with peers from another country. It also provided insight into the needs and working environment of fellow teachers in another Latin American country facing a situation similar to the one currently experienced in Mexico.

This shows that the global community is within reach, even to a public high school like the ENMSL. There is still a lot of work to be done, but as a first attempt it can be considered that the objective has been achieved since the students said they were satisfied with the results.

Now, the ENMSL is ready to continue with the COIL methodology and undertake a new project that will improve practices and continue innovating educationally.

Although COIL has been in existence for 20 years, it was the first time the COIL methodology was implemented in the entire history of the ENMSL.

## Acknowledgements

To all the students who collaborated with great enthusiasm on this ENMSL's first COIL project. To the principal of the ENMSL, Mr. Benjamin Chavez for the facilities provided to the authors to carry out this project. To our beloved relatives for all their support.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Copilot, an AI virtual assistant, was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The author further declares that, apart from Copilot, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Betancourt, E., Nicasio, D. & Ramos, R. (2026). La Escuela Preparatoria de Leon: Patrimonio de la Universidad de Guanajuato [The high school of Leon: University of Guanajuato Heritage]. 15. University of Guanajuato.

Hackett, S., Dawson, M., Janssen, J. & van Tartwijk1, J. (2024). Defining Collaborative Online International Learning (COIL) and Distinguishing it from Virtual Exchange. *TechTrends,* 68, 1078–1094.

State University of New York (2025). SUNY COIL. https://coil.suny.edu/

University of Guanajuato. (2025). Inician actividades académicas y culturales entre el CNMS y el Colegio Agustiniano Suba, de Bogotá. [Academic and cultural activities begin between the CNMS and the Colegio Agustiniano Suba in Bogota.] https://ugto.mx/noticias/quehacer-institucional/20287-inician-actividades-academicas-y-culturales-entre-el-cnms-y-el-colegio-agustiniano-suba-de-bogota

**Contact email:** jb.ortiz@ugto.mx

# From Majlis to Manuscript: Leveraging Emirati Oral Traditions to Scaffold Phase 1 Literacy in the PYP

Nguyilan Shaaji, Emirates National Schools, United Arab Emirates
Alueshima Jennifer Shaapera, Blossom Early Education Institution, United Arab Emirates

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Emirati children arrive in KG2 as competent storytellers. Most have spent their first five years in households where narrative is an ordinary daily practice: a grandmother recounting life before the oil economy, moral tales told at bedtime, the conversational turn-taking of the majlis. When formal English literacy instruction begins, it usually begins with imported texts about snow, forests and farmyard animals, and the confident five-year-old narrator often becomes a quiet one. This paper reports a small-scale practitioner action research study conducted in a single KG2 classroom in an English-medium IB Primary Years Programme (PYP) setting in the United Arab Emirates. Over several inquiry cycles, the class teacher and a research partner introduced three linked practices: co-constructed Storytelling Adventures built from Emirati folktales and family narratives; Word Detective play, in which children hunted for sounds and rhymes inside stories they already knew; and Visible Thinking routines anchored by familiar cultural artefacts. Data was generated through pedagogical documentation, audio-recorded storytelling sessions, anecdotal records, samples of child talk during play, and teacher reflective journals. These are observations from one classroom, generated by the teachers who taught it, and they carry the limits that implies. They do not establish that culturally familiar storytelling causes measurable literacy gains. What they do suggest is more modest and, we think, still worth saying: cultural familiarity appears to function as a scaffold for early English literacy rather than a decorative addition to it and treating it as such costs a teacher very little.

*Keywords:* early literacy, oral tradition, Emirati children, culturally sustaining pedagogy, IB Primary Years Programme, translanguaging, action research

iafor
The International Academic Forum
www.iafor.org

## Introduction

The claim that oral language precedes and supports written language is not new, and it is not seriously contested. Children learn to mean before they learn to read, through interaction, joint attention and shared narrative (Vygotsky, 1978; Wells, 1986). Longitudinal evidence has since put numbers to the intuition: narrative retelling and linguistic comprehension measured at age five make unique contributions to reading comprehension five and nine years later, even after controlling for cognitive ability, phonological skill and maternal education (Babayiğit et al., 2021), and preschool language skills sit on the main pathway to later comprehension in meta-analytic reviews of the field (Hjetland et al., 2020).

What is less often examined is whose oral language is being counted. The narrative skills a child brings to school are culturally shaped: they are stories of particular kinds, told in particular settings, about particular things. Emirati children typically arrive at kindergarten with a great deal of narrative competence developed in Arabic through family gatherings and the *majlis* which is a gathering space that UNESCO (2015) inscribed on the Representative List of the Intangible Cultural Heritage of Humanity specifically for its role in transmitting oral heritage, including folk stories, folk songs and Nabati poetry, largely to children who accompany adults there. This is what Moll et al. (1992) would call a fund of knowledge: a household resource that is real, transferable and generally invisible to the school.

The problem this study grew out of is a practical one. In an English-medium PYP classroom, the entry points to formal literacy are usually commercial schemes and picture books produced elsewhere. A five-year-old Emirati child meeting these texts is asked to do two unfamiliar things simultaneously; to decode a new language and infer a new cultural world with the second demand rarely acknowledged as a demand at all. Teachers notice the result before they can name it: children who talk fluently in the playground go silent on the carpet.

We wanted to know what would happen if we stopped treating the children's existing narrative repertoire as background and started treating it as curriculum. This paper describes what we did, what we saw, and what we think it does and does not show.

## The Emirati Early Years Context

The UAE is an unusually complex setting for early language education, and some of that complexity bears directly on this study. Emirati citizens are a minority within a superdiverse expatriate population, Arabic is the official language, and English functions as the working lingua franca and, in much of the private sector, as the medium of instruction from the earliest years (Gallagher, 2022). Successive policy efforts in Abu Dhabi have pursued Arabic–English bilingualism from kindergarten onwards, but the ambition of “balanced bilingualism” has proved considerably easier to state than to deliver (Gallagher, 2011, 2022).

Two features of this landscape shaped our thinking. The first is diglossia. Emirati children grow up speaking a spoken vernacular and encounter Modern Standard Arabic largely through school, television and books; the two varieties differ enough that the distance between them measurably affects children's phonological representations and their early reading in Arabic (Saiegh-Haddad, 2003; Saiegh-Haddad & Spolsky, 2014). Some scholars working in the UAE argue that the situation is better described as triglossia once the classroom presence of English is included (Carroll et al., 2017). The practical implication is that an Emirati five-year-old is not moving from one settled language into a second one. They are already negotiating varieties,

and English is a third demand layered onto an existing negotiation. Any account of "the child's home language" that treats Arabic as a single undifferentiated block will misdescribe what these children are actually doing.

The second feature is what happens at home. Barza and von Suchodoletz (2016) found that Emirati parents' preferences around shared reading function as a form of cultural transmission rather than simply as skills practice—the reading matters partly because of what it carries. Tibi and McLeod (2014) document how young children's Arabic language and literacy develops in the UAE specifically, and Al-Qaryouti et al. (2016) survey emergent literacy provision across the Gulf Cooperation Council states. Read together, this work suggests that the home is doing substantial language and literacy work that the English-medium classroom is not systematically connected to. Our study is, in one sense, an attempt to build a short bridge across that gap.

## Review of Relevant Literature

### Oral Language and Narrative as Literacy Foundations

Strong oral language underpins phonological awareness, vocabulary breadth and later reading comprehension (Hjetland et al., 2020; Snow et al., 1998). For bilingual learners, competence developed in a first language does not sit inertly alongside the second; it can support it, and the relationship between the two is developmental rather than zero-sum (Cummins, 2001; Hammer et al., 2014). Narrative is a particularly productive site for this because it demands extended, decontextualised talk of the kind that written language later requires. When children narrate, retell and argue about stories, they rehearse the structures they will meet again on the page (Wells, 1986).

### Familiarity, Cognitive Load and Willingness to Speak

Isbell et al. (2004) found that storytelling and story reading produced measurable differences in young children's oral language complexity and story comprehension. The mechanism most commonly proposed is a reduction in competing demand: when a child already knows how the story goes, attention is freed for how it sounds. Krashen's (1982) affective filter offers a complementary account at the emotional level—anxiety, low confidence and fear of error suppress production regardless of underlying competence. Both accounts predict the same thing in practice, which is that familiar content should increase the volume and the riskiness of a young learner's talk. That prediction is straightforwardly testable in a classroom, which is what drew us to it.

### Funds of Knowledge and Culturally Sustaining Pedagogy

Moll et al. (1992) and, later, González et al. (2005) reframed household knowledge from a deficit to be compensated for into an asset to be drawn on. Ladson-Billings (1995) made the pedagogical case that children learn more when their cultural backgrounds are treated as resources; Paris and Alim (2017) pushed further, arguing that relevance is not enough and that schooling should actively sustain the linguistic and cultural practices children arrive with, rather than using them as a stepping stone and then discarding them. The distinction matters for our design. Using an Emirati folktale as a warm-up before returning to the "real" scheme would be cultural relevance. Treating the folktale as the site where phonological and vocabulary work actually happens is closer to what Paris and Alim have in mind.

### Translanguaging in Early Years Classrooms

The translanguaging literature provides a way of describing what bilingual children do without treating their languages as separate compartments to be policed (García & Kleyn, 2016). Flynn et al. (2021) offer a close case study of an emergent bilingual preschooler whose storytelling improved in both of his languages when he was permitted to draw on his whole repertoire in the telling—an important finding, because the usual anxiety about translanguaging is that it will slow the acquisition of the target language. In the UAE specifically, Al-Bataineh and Gallagher (2021) examined how prospective teachers view the meshing of Arabic and English in children's storybooks, and found attitudes considerably more mixed than policy rhetoric might suggest. We took this as a caution as much as an endorsement: translanguaging in an English-medium school is a pedagogical choice teachers have to be able to defend.

### Play, Inquiry and Literacy in the PYP

The PYP treats literacy as transdisciplinary and socially constructed, emerging through inquiry and authentic experience rather than through isolated skill drills (International Baccalaureate Organization, 2018). This is philosophically compatible with what we wanted to do, though compatibility is not the same as ease. Pyle and Danniels (2017) describe a continuum of play-based learning and are candid about the risk at its centre: adult intentions can hijack play until it stops being play. Baker (2015) documents exactly this tension in Abu Dhabi kindergartens, where English-medium teachers found their beliefs about play difficult to enact under curricular and assessment pressure. We designed with that warning in view, and we return to it in our findings.

## Methodology

### Design and Positionality

This was a qualitative practitioner action research study, structured as iterative plan–act–observe–reflect cycles (Kemmis et al., 2014). We chose action research because the question was a teacher's question and the answer had to be usable on a Tuesday morning. It carries an obvious cost: we were not neutral observers of these children. One of us taught them daily and had a stake in the interventions working. We have tried to manage this by documenting what did not work alongside what did, by keeping reflective journals that record disconfirming as well as confirming instances, and by having the second author, who did not teach the class, review the documentation independently. Readers should nonetheless weigh the findings as those of interested parties.

### Setting and Participants

The study took place in one KG2 classroom in an English-medium school in the UAE following the IB Primary Years Programme. Participants were *32* Emirati children aged five, all English language learners, all with Arabic as the dominant home language and Emirati vernacular Arabic as their everyday spoken variety. The classroom operated on play-based, inquiry-driven principles consistent with PYP early years practice. The study ran across *3 terms per year.*

## The Three Interventions

### *Storytelling Adventures*

Children co-constructed English storytelling experiences drawn from Emirati folktales, desert and coastal animals, family gatherings and the moral tales commonly told at home. The design principle was to hold the narrative structure constant—the shapes children already knew—while varying the language. Children were not required to tell the story in English; they were required to help build it. In practice this meant a good deal of translanguaging, with Arabic used to negotiate meaning and English used to perform it, which is close to what García and Kleyn (2016) describe.

### *Word Detective Play*

During play-based centres, children hunted for sounds, rhymes and repeated words inside the stories they had just built. Framing this as detection rather than testing was deliberate: a detective who is wrong is still doing detective work, whereas a child who is wrong on the carpet is often just wrong. Target vocabulary was selected using Beck et al.'s (2013) Tier 2 criterion—high-utility words that recur across contexts—rather than the concrete nouns that dominate imported schemes.

### *Visible Thinking Routines*

We used *See–Think–Wonder* and related routines (Ritchhart et al., 2011) with authentic artefacts: the *dallah* coffee pot, traditional clothing, photographs of desert and coastal life, and prompts inviting children to bring a family story. The artefacts gave abstract English something physical to attach to, and—more usefully than we expected—gave children who were reluctant to speak something to point at.

## Data Generation

Data sources were pedagogical documentation of child talk in the Reggio-influenced sense of documentation as a shared, interpretable trace of learning rather than a record of outcomes (Rinaldi, 2006); anecdotal observational records; audio recordings of storytelling sessions; samples of children's language use during play; and teacher reflective journals maintained across all cycles.

## Analysis

Audio was transcribed and analysed inductively alongside documentation and journals using reflexive thematic analysis (Braun & Clarke, 2006). Coding proceeded in two passes: an initial open pass across the full corpus, then a focused pass in which candidate themes were tested against the whole dataset, including instances that did not fit. Both authors coded independently before comparing; disagreements were resolved by returning to the recordings rather than by averaging. Trustworthiness was pursued through prolonged engagement, triangulation across data types, and an audit trail of coding decisions (Lincoln & Guba, 1985). We make no claim to inter-rater reliability in a statistical sense, which would be inappropriate to this design.

## Ethics

Informed consent was obtained from parents and school leadership, and assent was sought from the children in age-appropriate terms before recording. Children could decline to participate in any activity or recording without consequence, and two occasions on which children did so are noted in the documentation. All names used in reporting are pseudonyms, and identifying details of families' stories have been altered where they could identify a household.

# Findings

## Willingness to Speak Arrived Before Accuracy Did

The clearest and earliest change was in volunteering. Children who had participated in English sessions only when directly addressed began putting their hands up, interrupting, and correcting the teacher's version of stories they knew better than she did. Crucially, this happened before any improvement in accuracy. The English was often wrong—mixed word order, Arabic function words carrying English content words, verbs left uninflected—but there was much more of it, and it was offered rather than extracted.

This is consistent with the affective account (Krashen, 1982) and with Isbell et al.'s (2004) findings on storytelling and oral language. We would add one observation from the documentation that we did not anticipate: the children who changed most were not the ones with the weakest English. They were the ones whose families told the most stories at home. Several children with comparatively strong English but less narrative-rich home environments changed very little. We cannot substantiate this pattern with the data we have, and we report it as a hypothesis worth testing properly rather than as a finding.

## Sound Became Easier to Attend to When Meaning Was Already Secure

During Word Detective play, children sustained phonological work—initial sound isolation, rhyme detection, syllable clapping—for noticeably longer when the words came from a story they had helped build than when the same skills were practised on scheme vocabulary. Sessions that had previously fragmented after a few minutes ran to their natural end.

The mechanism we infer is the one the literature proposes: comprehension was not competing for attention, so attention was available for form (Snow et al., 1998). There is a diglossic wrinkle worth naming. Some of the sound work children found easiest in English involved phonemes with close analogues in their spoken Emirati Arabic, and some of the hardest involved contrasts absent from it. This mirrors what Saiegh-Haddad (2003) found within Arabic itself, where distance from the spoken vernacular predicts difficulty. A properly designed study could examine whether the vernacular, rather than Modern Standard Arabic, is the more useful reference point for planning English phonics for these children. Our data can only raise the question.

## Children Brought the Home Into the Classroom Without Being Asked

By the *20th* cycle, children were arriving with stories. They had asked at home, been told something, and wanted to tell it. Some of these were folktales; more of them were family history—a grandfather's work, a relative's journey, a house that no longer exists. Children

negotiated these in Arabic with each other and then worked, sometimes laboriously, at rendering them in English for the teacher.

This is the finding we consider most significant, because it changed the direction of supply. The classroom stopped being the only place where legitimate content originated. That is precisely the shift Moll et al. (1992) and González et al. (2005) describe, and it also produced the translanguaging pattern Flynn et al. (2021) document, in which drawing on the full repertoire supported rather than displaced the developing second language. It also had a practical benefit we had not planned for: parents who had previously engaged with the school mainly around behaviour began asking what stories were needed next.

### What Did Not Work

Two things went less well, and reporting them seems more useful than not.

First, we hijacked the play. In the middle cycles, our enthusiasm for phonological outcomes turned *Word Detective* into a thinly disguised assessment. Children's talk dropped off immediately and did not recover until we pulled back and let the activity be genuinely open-ended again. This is exactly the failure mode Pyle and Danniels (2017) describe, and the fact that we had read the paper beforehand did not prevent us from committing it. Baker (2015) reports the same pressure operating on English-medium teachers in Abu Dhabi kindergartens, which suggests it is structural rather than personal.

Second, artefact-based routines were uneven. The dallah and the clothing generated sustained talk. Photographs of desert life generated much less, and on reflection we think we had chosen images that represented a heritage version of Emirati life rather than the lives these particular children actually lead in a city. Cultural familiarity is not the same as cultural iconography, and children can tell the difference faster than adults can.

## Discussion

Taken together, the observations point to a fairly narrow claim. Cultural familiarity in the narrative frame appears to lower the cost of participation for these young Emirati learners, and lowering that cost makes more language available to work on. This is a scaffolding claim, not a motivational one. The argument is not that children tried harder because they felt good; it is that a story they already possessed carried part of the cognitive load, and what it carried was released for other work.

Framed this way, the findings sit comfortably with Cummins (2001) on the relationship between identity affirmation and academic engagement, and with the culturally sustaining position that children's existing competencies are the material of instruction rather than a preamble to it (Paris & Alim, 2017). They also fit the PYP's own commitment to inquiry, agency and learner voice (International Baccalaureate Organization, 2018), though we would note that the fit is easier to describe in a paper than to sustain in a term, as section 5.4 indicates.

There is a broader point about the UAE context. Policy discussion of Arabic–English bilingualism has tended to concern medium of instruction, timing and balance (Carroll et al., 2017; Gallagher, 2011, 2022). Less attention has been paid to content—to what the English-medium classroom is actually about. Our small study suggests that the content question may be doing more work than it is credited with. A child can be in an early-immersion programme

with excellent structural design and still meet English exclusively through material that has nothing to do with them.

## Limitations

The limitations here are substantial and we would rather state them plainly than let a reader discover them.

This is one classroom, one cohort, one school type, and a small number of children. Nothing here generalises statistically, and it is not intended to. The teacher-researchers were personally invested in the interventions succeeding, and while we took the steps described in section 4.1, we cannot eliminate that. There was no comparison group and no pre- and post-testing, so claims about vocabulary retention or phonological gain cannot be made on this evidence; where an earlier version of this work described such gains as measurable, that was an overstatement, and we have corrected it. What we have is documented change in participation and talk, interpreted by the people who produced it.

There is also a selection issue in what got documented. Teachers notice children who start talking. Children who remained quiet throughout are underrepresented in our data, and we do not know whether the interventions did nothing for them or whether we simply failed to record it. A study designed to look for non-responders would be a more useful next step than a larger version of this one.

## Implications for Practice

For teachers in comparable settings, four things seem worth carrying forward.

Treat home oral traditions as instructional material, not as warm-up. The distinction is whether the phonological and vocabulary work happens inside the familiar story or after it.

Plan English phonics with the child's spoken variety in mind rather than a generic notion of “Arabic”. The vernacular is what the child actually has, and the distance between it and any target is what predicts difficulty (Saiegh-Haddad, 2003).

Ask families for stories rather than for reading minutes. The request is culturally legible in a way that a reading log often is not, and it recruits a resource that already exists (Barza & von Suchodoletz, 2016).

Protect the play. If the activity becomes assessment, the talk stops, and the talk was the point (Pyle & Danniels, 2017).

## Conclusion

We began with a teacher's observation that confident five-year-old storytellers were becoming silent English learners, and we asked whether the silence was about language or about content. On the evidence of one classroom, some of it was about content. When these children met English inside narratives they already owned, they talked more, risked more, and brought more of their households into the room with them.

We would not want to make this sound larger than it is. A single KG2 class, a handful of inquiry cycles, and two invested teachers do not settle anything. But the intervention was cheap, it required no new resources beyond a coffee pot and a willingness to ask, and it moved the thing early years teachers most want to move, which is how much a child is prepared to say. If that is right, then a good deal of the literacy resource these children need is already in the house. The *majlis* has been doing this work for a long time. The classroom has mostly not been listening.

## References

Al-Bataineh, A., & Gallagher, K. (2021). Attitudes towards translanguaging: How future teachers perceive the meshing of Arabic and English in children's storybooks. *International Journal of Bilingual Education and Bilingualism, 24*(3), 386–400.

Al-Qaryouti, I., Ihmeideh, F., Al Bustami, G., & Homidi, M. (2016). Evidence-based strategies to support children's emergent literacy in the Gulf Cooperation Council (GCC) countries. *European Journal of Developmental Psychology, 13*(5), 545–562.

Babayiğit, S., Roulstone, S., & Wren, Y. (2021). Linguistic comprehension and narrative skills predict reading ability: A 9-year longitudinal study. *British Journal of Educational Psychology, 91*(1), 148–168.

Baker, F. (2015). Challenges presented to personal theories, beliefs and practices of play in Abu Dhabi kindergartens: The English medium teacher perspective. *Early Years, 35*(1), 22–35.

Barza, L., & von Suchodoletz, A. (2016). Home literacy as cultural transmission: Parent preferences for shared reading in the United Arab Emirates. *Learning, Culture and Social Interaction, 11*, 142–152.

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). Guilford Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Carroll, K., Al Kahwaji, B., & Litz, D. (2017). Triglossia and promoting Arabic literacy in the United Arab Emirates. *Language, Culture and Curriculum, 30*(3), 317–332.

Cummins, J. (2001). *Negotiating identities: Education for empowerment in a diverse society* (2nd ed.). California Association for Bilingual Education.

Flynn, E. E., Hoy, S. L., Lea, J. L., & García, M. A. (2021). Translanguaging through story: Empowering children to use their full language repertoire. *Journal of Early Childhood Literacy, 21*(2), 283–309.

Gallagher, K. (2011). Bilingual education in the UAE: Factors, variables and critical questions. *Education, Business and Society: Contemporary Middle Eastern Issues, 4*(1), 62–79.

Gallagher, K. (2022). Early language education in the United Arab Emirates. In M. Schwartz (Ed.), *Handbook of early language education* (pp. 893–921). Springer.

García, O., & Kleyn, T. (Eds.). (2016). *Translanguaging with multilingual students: Learning from classroom moments*. Routledge.

González, N., Moll, L. C., & Amanti, C. (Eds.). (2005). *Funds of knowledge: Theorizing practices in households, communities, and classrooms*. Lawrence Erlbaum.

Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C., & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly, 29*(4), 715–733.

Hjetland, H. N., Brinchmann, E. I., Scherer, R., Hulme, C., & Melby-Lervåg, M. (2020). Preschool pathways to reading comprehension: A systematic meta-analytic review. *Educational Research Review, 30*, 100323.

International Baccalaureate Organization. (2018). *Learning and teaching*. IB Publishing.

Isbell, R., Sobol, J., Lindauer, L., & Lowrance, A. (2004). The effects of storytelling and story reading on the oral language complexity and story comprehension of young children. *Early Childhood Education Journal, 32*(3), 157–163.

Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The action research planner: Doing critical participatory action research*. Springer.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465–491.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Moll, L. C., Amanti, C., Neff, D., & González, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice, 31*(2), 132–141.

Paris, D., & Alim, H. S. (Eds.). (2017). *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world*. Teachers College Press.

Pyle, A., & Danniels, E. (2017). A continuum of play-based learning: The role of the teacher in play-based pedagogy and the fear of hijacking play. *Early Education and Development, 28*(3), 274–289.

Rinaldi, C. (2006). *In dialogue with Reggio Emilia: Listening, researching and learning*. Routledge.

Ritchhart, R., Church, M., & Morrison, K. (2011). *Making thinking visible: How to promote engagement, understanding, and independence for all learners*. Jossey-Bass.

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics, 24*(3), 431–451.

Saiegh-Haddad, E., & Spolsky, B. (2014). Acquiring literacy in a diglossic context: Problems and prospects. In E. Saiegh-Haddad & R. M. Joshi (Eds.), *Handbook of Arabic literacy* (pp. 225–240). Springer.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. National Academy Press.

Tibi, S., & McLeod, L. (2014). The development of young children's Arabic language and literacy in the United Arab Emirates. In E. Saiegh-Haddad & R. M. Joshi (Eds.), *Handbook of Arabic literacy* (pp. 303–321). Springer.

UNESCO. (2015). *Majlis, a cultural and social space*. Representative List of the Intangible Cultural Heritage of Humanity. https://ich.unesco.org/en/RL/majlis-a-cultural-and-social-space-01076

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wells, G. (1986). *The meaning makers: Children learning language and using language to learn*. Heinemann.

**Contact email:** nguyilanshaaji@yahoo.com

# Total Quality Management in Action: What Headmasters of High School Believe

Vasileios Tsiantos, Democritus University of Thrace, Greece
Argiris Fyntanakis, Democritus University of Thrace, Greece

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This paper examines Total Quality Management (TQM) as a management philosophy that can support improvement in Greek secondary schools, with emphasis on the role of the principal. The aim is to record and interpret school leaders' knowledge and understanding of TQM principles within the framework of institutionalised internal school self-evaluation and alignment with European and international quality policies. Specifically, the research investigates whether knowledge and understanding of TQM principles differ according to gender, years of service in a school leadership position, TQM training, and postgraduate studies in educational administration; develops a predictive model; and identifies adoption profiles through cluster analysis. The study employed a quantitative approach using a structured five-point Likert-scale questionnaire, distributed electronically to principals in the Regional Directorates of Education of Central Macedonia and Eastern Macedonia–Thrace (N = 314). The instrument's factor structure was examined through principal component analysis with Varimax rotation, while reliability was assessed using Cronbach's alpha. Group comparisons were conducted using non-parametric tests, correlations with Spearman's coefficient, and predictive analysis through multiple linear regression. Women and principals with administration-related postgraduate studies reported higher levels of TQM knowledge and understanding, whereas differences by leadership tenure were not statistically significant. TQM training emerged as the strongest predictor, and the regression model explained approximately 29% of the variance. Positive correlations were found among the four dimensions, while cluster analysis identified three distinct adoption profiles. The findings support targeted professional development aimed at strengthening principals' readiness for quality-oriented school improvement.

*Keywords:* total quality management, school self-evaluation, secondary school principals, quality improvement, Greece

iafor
The International Academic Forum
www.iafor.org

## Introduction

The current Greek framework defines school self-evaluation as a continuous and participatory process linked to annual collective planning, improvement actions, implementation, and appraisal. Within this framework, secondary school principals coordinate processes through which schools identify priorities, set objectives, monitor collective actions, and use evidence to support improvement (Institute of Educational Policy, 2024; Ministry of Education and Religious Affairs, 2021). At the European and international levels, policy documents similarly emphasise the role of school leadership and the integration of evaluation and assessment into coherent frameworks for educational improvement (Council of the European Union, 2020; OECD, 2013).

Within this institutional context, Total Quality Management (TQM) can be understood as a management philosophy centred on continuous improvement, stakeholder participation, collective responsibility, leadership commitment, and evidence-informed decision-making (Bouranta et al., 2021; Sfakianaki, 2019; Sfakianaki et al., 2021). In education, its application involves aligning organisational processes with school improvement goals, supporting teacher participation and collaboration, and responding to the needs of students and other school stakeholders (Glaveli et al., 2022; Saiti, 2012; Sfakianaki et al., 2023). School self-evaluation and TQM share several operational features, including collective goal-setting, systematic feedback, the use of evidence, and continuous improvement (Institute of Educational Policy, 2024; Ministry of Education and Religious Affairs, 2021; Sfakianaki, 2019). In the present study, this correspondence provides the basis for examining school self-evaluation as an institutional process through which quality-oriented practices may be developed at school level.

Previous Greek studies have developed instruments for examining TQM in primary and secondary education and have investigated teachers' perceptions and its implementation in primary education and other specific educational settings (Kaiseroglou et al., 2023, 2025; Sfakianaki, 2019; Sfakianaki et al., 2023). Within the literature reviewed for the present study, evidence focused specifically on secondary school principals remains limited, particularly regarding their understanding of TQM principles and the way they relate these principles to institutionalised school self-evaluation and the wider European and international quality agenda (Bouranta et al., 2021; Kaiseroglou et al., 2023, 2025; Sfakianaki, 2019; Sfakianaki et al., 2023). This issue is important because principals coordinate the school-level processes through which quality policies are interpreted and implemented (Institute of Educational Policy, 2024; Ministry of Education and Religious Affairs, 2021; Saiti, 2012).

The present study examines Greek secondary school principals' knowledge and understanding of TQM principles. It investigates differences according to gender, leadership tenure, TQM-related training, and administration-related postgraduate studies. It also examines the combined predictive contribution of these characteristics, the relationships among knowledge and understanding, school self-evaluation, quality improvement, and European and international harmonisation, and the existence of distinct profiles across these four dimensions. The study seeks to provide evidence that can inform targeted professional development and organisational support for secondary school principals.

# Theoretical Background

## Total Quality Management and School Leadership

Total Quality Management is an integrated management philosophy centred on continuous improvement, collective responsibility, stakeholder participation, and the systematic use of evidence (Bouranta et al., 2021; Sfakianaki, 2019; Sfakianaki et al., 2021). Its application in education concerns the quality of organisational and educational processes and the capacity of a school to identify needs, evaluate its work, and pursue improvement over time (Glaveli et al., 2022; Saiti, 2012; Sfakianaki et al., 2023). In school settings, TQM is associated with quality-oriented leadership, the active participation and empowerment of teachers, attention to the needs of students, parents, and the local community, and evidence-informed decision-making (Bouranta et al., 2021; Saiti, 2012; Sfakianaki et al., 2021).

Within this framework, the principal coordinates processes through which shared goals are established, responsibilities are distributed, improvement actions are monitored, and evidence is incorporated into decision-making (Saiti, 2012; Sfakianaki et al., 2021). Leadership therefore contributes to the development of the organisational conditions required for collaboration, professional learning, and sustained improvement (Bouranta et al., 2021; Glaveli et al., 2022). In the present study, knowledge and understanding of TQM refers to principals' ability to recognise and interpret these principles and relate them to the everyday operation of the school.

## School Self-Evaluation in Greece

The current Greek framework defines internal school evaluation as a collective, participatory, and evidence-informed process. The teaching staff participate in the appraisal of the school's educational work, identify priorities, and contribute to working groups organised around specific axes and indicators (Institute of Educational Policy, 2024; Ministry of Education and Religious Affairs, 2021). The process follows an annual cycle that includes appraisal, planning, implementation, monitoring, and feedback. Qualitative and quantitative evidence is used to identify strengths and areas requiring improvement and to support the design and evaluation of action plans (Institute of Educational Policy, 2024; Ministry of Education and Religious Affairs, 2021).

This institutional process shares several operational principles with TQM, including collective responsibility, participation, systematic feedback, evidence use, and continuous improvement (Bouranta et al., 2021; Institute of Educational Policy, 2024; Ministry of Education and Religious Affairs, 2021; Sfakianaki, 2019). On this basis, school self-evaluation can be regarded as an organisational setting within which quality-oriented practices may be developed. The existence of a formal process does not by itself determine how it is understood or enacted within individual schools, since the implementation of quality-management principles depends on leadership, staff involvement, organisational culture, and institutional support (Bouranta et al., 2021; Sfakianaki et al., 2021). Principals' knowledge, interpretations, and leadership practices influence how collective planning, evidence use, and improvement actions are coordinated at school level (Institute of Educational Policy, 2024; Saiti, 2012; Sfakianaki et al., 2021).

### European and International Policy Context

The Greek framework can also be situated within a broader policy context concerned with educational quality, evaluation, and school improvement. The frameworks considered in the present study include the Strategic Framework for European Cooperation in Education and Training, the Council Recommendation on Key Competences for Lifelong Learning, the United Nations 2030 Agenda for Sustainable Development, and the OECD's *Synergies for Better Learning* (Council of the European Union, 2009, 2018; OECD, 2013; United Nations, 2015).

These documents differ in purpose, scope, and institutional status. The European and United Nations frameworks emphasise educational quality, participation, learner-centred education, inclusion, and continuing improvement (Council of the European Union, 2009, 2018; United Nations, 2015). The OECD framework places particular emphasis on the systematic use of evaluation and assessment evidence to improve educational processes, accountability, and decision-making (OECD, 2013). The Greek self-evaluation framework similarly incorporates collective planning, participation, monitoring, evidence use, feedback, and improvement actions (Institute of Educational Policy, 2024; Ministry of Education and Religious Affairs, 2021). These features also correspond to principles identified in educational applications of TQM, particularly continuous improvement, participation, collective responsibility, leadership commitment, and evidence-informed decision-making (Bouranta et al., 2021; Sfakianaki, 2019; Sfakianaki et al., 2021).

In the present study, European and international harmonisation refers to principals' perceptions of the correspondence between school-level practices and these wider policy orientations. This dimension provides a basis for examining whether principals connect internal evaluation and quality improvement with the broader institutional environment in which Greek schools operate.

## Methodology

### Research Design and Participants

The study employed a quantitative survey design using a structured questionnaire. The questionnaire was distributed electronically via Microsoft Forms to 765 principals of secondary schools in the Regional Directorates of Education of Central Macedonia and Eastern Macedonia–Thrace. The email addresses were obtained from the central registry of school units maintained by the Greek School Network. A total of 314 completed questionnaires were returned, yielding a response rate of 41.05%. The sample comprised 158 men (50.3%) and 156 women (49.7%). Most participants had extensive professional experience, with 88.6% reporting more than 20 years of service in public education.

### Ethical Considerations

Ethical approval was granted by the Research Ethics Committee of Democritus University of Thrace at its 9th meeting on 20 May 2025 (protocol no. ΔΠΘ/ΕΗΔΕ/62741/530, issued 22 May 2025). Participation was voluntary, and participants were informed about the purpose of the study, the use of the collected data, and their right to withdraw. Informed consent was obtained before completion of the questionnaire. Responses were collected anonymously and handled confidentially.

## Questionnaire and Measurement Model

The questionnaire was developed by adapting items from established Greek and international instruments examining TQM in education and was adjusted to the context of Greek secondary schools (Glaveli et al., 2022; Papadopoulos et al., 2022; Sfakianaki, 2019). Responses were recorded on five-point Likert-type scales, with anchors ranging from 1 to 5 according to the wording of each item.

An exploratory factor analysis was conducted using principal component analysis with Varimax rotation. The final 19-item solution demonstrated adequate sampling suitability, with a Kaiser–Meyer–Olkin value of .890 and a statistically significant Bartlett's test of sphericity, $\chi^2(171) = 3821.329$, $p < .001$. Four factors were retained, accounting for 68.149% of the total variance. As shown in Table 1, the four scales comprised between four and six items and demonstrated satisfactory to high internal consistency, with Cronbach's α coefficients ranging from .836 to .889.

**Table 1**
*Factor Structure and Internal Consistency of the Measurement Model*

| Factor | Number of items | Cronbach's α |
|---|---|---|
| Knowledge and understanding of TQM principles | 6 | .889 |
| School self-evaluation | 5 | .871 |
| Quality improvement | 4 | .836 |
| European and international harmonisation | 4 | .843 |

*Note.* TQM = Total Quality Management.

## Data Analysis

Distributional characteristics were examined using the Kolmogorov–Smirnov and Shapiro–Wilk tests, together with Q–Q and detrended Q–Q plots. An independent-samples t test with equal variances not assumed was used to examine gender differences. The Mann–Whitney U test was used to compare principals with and without TQM-related training, while the Kruskal–Wallis test was applied to comparisons involving leadership-tenure groups and levels of administration-related postgraduate study. Bonferroni-adjusted Mann–Whitney U post hoc comparisons were conducted following statistically significant Kruskal–Wallis results.

Spearman's rank correlation coefficient was used to examine associations among the four factor scores. Multiple linear regression was conducted with knowledge and understanding of TQM principles as the dependent variable and gender, leadership tenure, TQM-related training, and administration-related postgraduate studies as predictors. The independence of residuals and multicollinearity were examined using the Durbin–Watson statistic and collinearity diagnostics. Finally, cluster analysis was applied to the four factor scores to identify groups of principals with similar response profiles.

# Results

## Group Differences in TQM Knowledge and Understanding

The first set of analyses examined whether principals' knowledge and understanding of TQM principles differed according to demographic and professional characteristics. Women reported

higher knowledge and understanding scores ($M = 3.96$) than men ($M = 3.66$). An independent-samples $t$ test, with equal variances not assumed, showed that this difference was statistically significant, $t(296.936) = -3.098$, $p = .002$. The estimated effect size was small to moderate, Cohen's $d \approx 0.35$.

Differences among the four leadership-tenure groups did not reach statistical significance, $\chi^2(3) = 7.432$, $p = .059$. Although the highest scores were observed among principals with nine or more years of leadership experience, the evidence did not support a statistically significant difference among the tenure groups.

Principals who had participated in TQM-related training reported significantly higher knowledge and understanding scores than those who had not received such training, $p < .001$. Statistically significant differences were also found among levels of administration-related postgraduate study, $p < .001$. The highest mean score was recorded among principals with postgraduate studies in public administration or educational administration ($M = 4.18$).

## Predictive Model for TQM Knowledge and Understanding

Multiple linear regression was used to examine the combined contribution of gender, leadership tenure, TQM-related training, and administration-related postgraduate studies to principals' knowledge and understanding of TQM principles. The model was statistically significant, $F(4, 309) = 31.942$, $p < .001$, and explained 29.3% of the variance, $R^2 = .293$.

TQM-related training was the strongest standardised predictor, $\beta = .429$. Gender, $\beta = .190$, administration-related postgraduate studies, $\beta = .140$, and leadership tenure, $\beta = .103$, also made positive and statistically significant contributions to the model. The Durbin–Watson value was approximately 1.998, while variance inflation factors ranged from 1.03 to 1.14, indicating no substantial residual autocorrelation or multicollinearity.

## Correlations Among the Four TQM Dimensions

Spearman's rank correlations were positive and statistically significant for all pairs of factor scores. As shown in Table 2, the coefficients ranged from .313 to .550, with the strongest relationship observed between school self-evaluation and quality improvement.

**Table 2**
*Spearman Correlations Among the Four TQM Dimensions*

| Variable | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1. Knowledge and understanding of TQM principles | — | | | |
| 2. School self-evaluation | .313 | — | | |
| 3. Quality improvement | .448 | .550 | — | |
| 4. European and international harmonisation | .493 | .546 | .507 | — |

*Note.* $N = 314$. All correlations were statistically significant at $p < .01$.

## Adoption Profiles

Cluster analysis identified three groups of principals with different score patterns across the four factors. As shown in Table 3, Profile 3 was the largest group and displayed high mean scores across all four dimensions. Profile 1 combined moderate scores in knowledge and

understanding, quality improvement, and harmonisation with a lower score in school self-evaluation. Profile 2 showed the lowest scores in knowledge and understanding and harmonisation, while its scores for school self-evaluation and quality improvement were higher. The profiles therefore differed in both the overall level and the configuration of their scores across the four TQM-related dimensions.

**Table 3**
*Mean Factor Scores by Adoption Profile*

| Profile | *N* | Knowledge and understanding | School self-evaluation | Quality improvement | Harmonisation |
|---|---|---|---|---|---|
| Profile 1 | 65 | 3.77 | 2.82 | 3.82 | 3.25 |
| Profile 2 | 52 | 2.34 | 3.47 | 3.81 | 2.88 |
| Profile 3 | 197 | 4.20 | 4.11 | 4.65 | 4.00 |

*Note.* Values are mean factor scores.

## Discussion

### Professional Preparation and Knowledge of TQM

The clearest finding concerns the association between structured professional preparation and principals' knowledge and understanding of TQM. TQM-related training emerged as the strongest predictor, while administration-related postgraduate studies made an additional positive contribution. Leadership tenure contributed more weakly to the regression model and did not produce statistically significant differences across tenure groups. These findings indicate that principals' conceptual readiness for TQM is associated more strongly with targeted learning opportunities than with experience accumulated in a leadership position. This interpretation is consistent with research that identifies leadership development, staff participation, collaboration, and evidence-informed decision-making as important conditions for quality management in education (Bouranta et al., 2021; Saiti, 2012; Sfakianaki et al., 2021).

The findings do not establish that training or postgraduate study directly causes stronger TQM knowledge. Participation in such programmes may also be associated with prior interest in school leadership, quality management, or professional development. Nevertheless, the strength of the training coefficient, together with previous research on the conditions supporting TQM in schools, indicates that structured professional development should be treated as an important component of quality-oriented school leadership (Bouranta et al., 2021; Sfakianaki et al., 2021). The regression model explained 29.3% of the variance. The remaining variance is consistent with literature identifying organisational culture, staff involvement, leadership support, collaboration, and available resources as relevant conditions for the implementation of quality-management principles (Bouranta et al., 2021; Glaveli et al., 2022; Sfakianaki et al., 2021).

Women reported higher knowledge and understanding scores than men, with a small-to-moderate effect size. One possible interpretation is that this difference may be associated with leadership orientations that emphasise consultation, collaboration, participation, support, and staff empowerment. Coleman (2007) found that relational and collaborative descriptors were prominent in secondary headteachers' accounts of leadership, although such orientations were reported by both women and men and should not be treated as inherent gender characteristics. These orientations overlap with the soft dimensions of TQM, which emphasise teamwork,

communication, trust, participation, and employee involvement (Glaveli et al., 2022). The higher scores recorded by women in the present study may therefore reflect a closer reported alignment with relational and participatory principles associated with TQM. This interpretation remains tentative, since the questionnaire did not directly measure leadership style or actual leadership practices. Further research using broader samples and qualitative methods is needed to examine the processes underlying the observed gender difference.

## Relationships Among the Four Dimensions

The positive correlations among knowledge and understanding, school self-evaluation, quality improvement, and European and international harmonisation indicate that these dimensions tend to vary together. The strongest relationship was observed between school self-evaluation and quality improvement. This finding is consistent with the role assigned to self-evaluation in the Greek institutional framework, where schools identify priorities, design and monitor action plans, use evidence and feedback, and organise improvement processes (Institute of Educational Policy, 2024; Ministry of Education and Religious Affairs, 2021). It is also compatible with research linking school evaluation processes with quality improvement and TQM-oriented practices in education (Kaiseroglou et al., 2025; Sfakianaki, 2019).

The associations with European and international harmonisation suggest that principals who reported stronger alignment with wider quality-policy orientations also tended to report higher scores in school self-evaluation, quality improvement, and knowledge and understanding of TQM. The four dimensions therefore appear to represent related components of a broader quality orientation while retaining distinct meanings within the measurement model. Since the relationships are correlational, the findings do not show that stronger knowledge produces more favourable views of self-evaluation or that policy alignment leads to quality improvement. They show that higher and lower scores tend to coexist across the dimensions. Accordingly, the observed associations should not be interpreted as causal relationships.

## Interpreting the Adoption Profiles

The cluster analysis adds an important layer to the results by showing that principals' responses are organised into different configurations rather than along a single continuum of adoption. Profile 3, which included the majority of participants, displayed high scores across all four dimensions. It represents the configuration with the strongest alignment among TQM knowledge, school self-evaluation, quality improvement, and perceived policy harmonisation. These scores indicate favourable reported positions across the four dimensions, although they do not demonstrate that TQM principles are consistently implemented in school practice.

Profile 1 combined moderate scores in knowledge and understanding, quality improvement, and harmonisation with a distinctly lower score in school self-evaluation. This pattern may indicate a weaker connection between general support for quality improvement and perceptions of the institutional self-evaluation process. It could also reflect critical views of the current process or difficulties in relating its formal procedures to broader quality principles. The quantitative data cannot determine which of these explanations is more accurate.

Profile 2 showed the lowest scores in knowledge and understanding and in European and international harmonisation, while its scores for school self-evaluation and quality improvement remained moderate. This configuration may indicate that some principals engage positively with school improvement and self-evaluation without connecting these practices

strongly to TQM terminology or to the wider policy context. It should therefore not be interpreted as general resistance to quality or as evidence that self-evaluation is approached exclusively as bureaucratic compliance.

The three profiles should be treated as descriptive typologies derived from this sample. They require replication in independent samples before they can be considered stable categories of secondary school principals. Qualitative interviews could also clarify the meanings that participants assign to self-evaluation, quality improvement, and policy harmonisation within each profile.

### Implications for Professional Support

The differentiated profiles suggest that a uniform professional development programme may not respond equally well to all principals. Participants with limited knowledge and understanding may benefit from foundational work on TQM principles, evidence use, stakeholder participation, and continuous improvement. Principals who express support for quality improvement while reporting weaker positions towards self-evaluation may require support focused on the relationship between collective planning, action plans, feedback, and school development. Participants with consistently high scores may benefit from advanced professional learning and opportunities to share practices across schools.

These implications remain provisional because the study is cross-sectional, based on self-reported data, and limited to two Regional Directorates of Education. They nevertheless provide a more defensible basis for professional development than classifications based solely on years of leadership experience. The central finding is that knowledge, self-evaluation, improvement orientation, and policy alignment are related dimensions whose configuration differs substantially among principals.

## Conclusion

This study provides empirical evidence on Greek secondary school principals' knowledge and understanding of Total Quality Management within the institutional context of school self-evaluation and the wider European and international quality-policy environment. TQM-related training emerged as the strongest predictor of knowledge and understanding, while administration-related postgraduate studies made an additional positive contribution. Leadership tenure did not produce statistically significant differences across tenure groups, indicating that professional experience alone does not necessarily correspond to stronger conceptual understanding of TQM principles.

The positive associations among TQM knowledge and understanding, school self-evaluation, quality improvement, and European and international harmonisation indicate that these dimensions form related components of principals' broader quality orientation. The three profiles identified through cluster analysis further show that principals combine these dimensions in different ways. These profiles support the development of differentiated forms of professional learning and organisational support, while remaining descriptive classifications that require replication in independent samples.

Several limitations should be considered. The study relied on self-reported data and may therefore be affected by socially desirable responding. Its cross-sectional design precludes causal conclusions, and the regression model explained 29.3% of the variance, leaving a

substantial proportion associated with factors not examined in the study. The sample was also restricted to secondary school principals from two Regional Directorates of Education, limiting the generalisability of the findings. Future research could combine survey data with interviews, observations, and evidence from school improvement processes, use longitudinal designs, and examine whether the three profiles are reproduced in broader samples. Such research could clarify how professional preparation, organisational conditions, and institutional support relate to principals' quality-oriented practices over time.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that no AI or AI-assisted technologies have been used to generate, refine, or correct the content in the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Bouranta, N., Psomas, E., & Antony, J. (2021). Findings of quality management studies in primary and secondary education: A systematic literature review. *The TQM Journal*, *33*(3), 729–769. https://doi.org/10.1108/TQM-02-2020-0020

Coleman, M. (2007). Gender and educational leadership in England: A comparison of secondary headteachers' views over time. *School Leadership & Management*, *27*(4), 383–399. https://doi.org/10.1080/13632430701562991

Council of the European Union. (2009). Council conclusions of 12 May 2009 on a strategic framework for European cooperation in education and training ('ET 2020'). *Official Journal of the European Union*, *C 119*, 2–10.

Council of the European Union. (2018). Council Recommendation of 22 May 2018 on key competences for lifelong learning. *Official Journal of the European Union*, *C 189*, 1–13.

Council of the European Union. (2020). Council conclusions on European teachers and trainers for the future. *Official Journal of the European Union*, *C 193*, 11–19.

Glaveli, N., Vouzas, F., & Roumeliotou, M. (2022). The soft side of TQM and teachers' job satisfaction: An empirical investigation in primary and secondary education. *The TQM Journal*, *34*(5), 922–938. https://doi.org/10.1108/TQM-11-2020-0269

Institute of Educational Policy. (2024). *Syllogikos programmatismos, esoteriki kai exoteriki axiologisi tou ergou tis scholikis monadas: Odigos efarmogis [Collective planning, internal and external evaluation of the school unit's work: Implementation guide]*. https://www2.iep.edu.gr/images/axiologisi/odigos/2024-12-03_%CE%9F%CE%B4%CE%B7%CE%B3%CF%8C%CF%82_%CE%91%CE%BE%CE%B9%CE%BF%CE%BB%CF%8C%CE%B3._Sep._2024.pdf

Kaiseroglou, N., Sfakianaki, E., & Koemtzi, M. (2023). Impact of teachers' demographics on Total Quality Management parameters: The case of primary education. *Education Sciences*, *13*(7), 679. https://doi.org/10.3390/educsci13070679

Kaiseroglou, N., Sfakianaki, E., & Siontorou, C. (2025). Evaluation and total quality management: The case of primary education. *The TQM Journal*, *37*(7), 1992–2014. https://doi.org/10.1108/TQM-05-2024-0165

Ministry of Education and Religious Affairs. (2021). *Syllogikos programmatismos, esoteriki kai exoteriki axiologisi ton scholikon monadon os pros to ekpaideftiko tous ergo [Collective planning and internal and external evaluation of school units with regard to their educational work]*. National Printing Office. https://www.minedu.gov.gr/publications/docs2020/%CE%A6%CE%95%CE%9A_%CE%92_4189.pdf

OECD. (2013). *Synergies for better learning: An international perspective on evaluation and assessment*. OECD Publishing. https://doi.org/10.1787/9789264190658-en

Papadopoulos, A., Kouli, O., Tsitskari, E., & Grammatikopoulos, V. (2022). Metrisi Dioikisis Olikis Poiotitas stin Elliniki Ekpaidefsi: Mia pilotiki erevna [Measurement of Total Quality Management in Greek education: A pilot research]. *Inquiries in Physical Education and Sport*, *20*(1), 26–40. https://doi.org/10.26253/heal.uth.ojs.ispe.2022.1572

Saiti, A. (2012). Leadership and quality management: An analysis of three key features of the Greek education system. *Quality Assurance in Education*, *20*(2), 110–138. https://doi.org/10.1108/09684881211219370

Sfakianaki, E. (2019). A measurement instrument for implementing total quality management in Greek primary and secondary education. *International Journal of Educational Management*, *33*(5), 1065–1081. https://doi.org/10.1108/IJEM-08-2018-0245

Sfakianaki, E., Kaiseroglou, N., & Kakouris, A. (2023). An instrument for studying TQM implementation in primary education: Development and empirical investigation. *Quality Assurance in Education*, *31*(3), 452–468. https://doi.org/10.1108/QAE-10-2022-0189

Sfakianaki, E., Kakouris, A., & Siontorou, C. (2021). Critical success factors for total quality management in primary and secondary education. *International Journal of Services and Operations Management*, *40*(4), 564–595. https://doi.org/10.1504/IJSOM.2021.120061

United Nations. (2015). *Transforming our world: The 2030 Agenda for Sustainable Development*. https://docs.un.org/en/A/RES/70/1

# Overview of Industrial Engineering Courses in Brazil Under the Focus of Entrepreneurship

Adriana de Paula Lacerda Santos, Federal University of Parana, Brazil
Amanda Santos, Pontifical Catholic University of Parana, Brazil
Bruno Lacerda Santos, Federal University of Technology of Parana, Brazil

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This study presents an overview of Industrial Engineering programmes in Brazil with a focus on entrepreneurship. The research is grounded in knowledge management concepts, including a brief discussion of the knowledge society and organisational knowledge. The main objective was to analyse the curriculum of the Industrial Engineering programme at the Federal University of Paraná from an entrepreneurship perspective. To support this analysis, a national panorama of Industrial Engineering programmes was developed, identifying their main characteristics and examining how entrepreneurship education has been incorporated into higher education institutions. A descriptive research method was adopted, using data collected from the National Institute for Educational Studies and Research (INEP) database. The results indicate a strong concentration of programmes in the Southeast region, representing sixty-one per cent of the total. The study also identified regional differences in the distribution of entrepreneurship-related subjects and highlighted the significant number of recently created programmes. Out of 805 programmes analysed, 443 had not yet graduated their first cohort, corresponding to fifty-five per cent, which demonstrates the relatively recent expansion of Industrial Engineering education in Brazil. This rapid growth reflects the increasing demand for professionals in this field. The findings further reveal that the curriculum of the Federal University of Paraná remains comparatively traditional when contrasted with other Brazilian institutions. The study suggests the need for curricular revision, emphasising the inclusion of entrepreneurship-related subjects to ensure that future engineers develop entrepreneurial competencies alongside technical skills, thereby strengthening innovation and professional adaptability.

*Keywords: e*ntrepreneurship, production engineering education, higher education, curriculum analysis, engineering education, knowledge management

iafor
The International Academic Forum
www.iafor.org

## Introduction

This study builds upon the research paper titled "Proposal for the Inclusion of Entrepreneurship in the Production Engineering Curriculum," which aimed to meet the contemporary demand for professionals capable of proposing innovations across diverse societal spheres, focusing primarily on small enterprises and intrapreneurship (Pellin and Santos, 2008).

Knowledge management demonstrates how not only organisations but also individuals are placing increasing value on knowledge (intangible assets) and less on tangible goods. This shift stems from the critical importance of knowledge sharing and dissemination. Unlike financial resources—whereby transferring funds to another party results in a personal loss—knowledge multiplies when shared, akin to learning a new language. Furthermore, if a learner goes years without sharing this knowledge, it becomes obsolete. However, to grasp the core meaning of Knowledge Management, one must first understand the distinctions between data, information, and knowledge, which are elucidated in the subsequent pages of this article.

## Entrepreneurial Pedagogy

The entrepreneurial pedagogy proposed by Fernando Dolabela is a didactic strategy that utilises the Theory of Dreams to develop the entrepreneurial capacity of students in pre-school, primary, and secondary education. This theory initially proposes that students formulate a dream—a future state where they wish to arrive, be, or exist. In the second stage, students must endeavour to realise this dream, actively pursuing everything necessary to achieve this objective. This strategy is reintroduced at the commencement of each academic year throughout the 14 years of basic education. Simple language is maintained throughout this process, oriented around the core matrix: "What is your dream, and how are you going to realise it?" (Dolabela, 2003).

The supporting elements presented by Filion (1991a, 1991b) and adapted within entrepreneurial pedagogy aim to develop conditions for self-knowledge, skills to understand the reality in which the individual is embedded, and competences to comprehend the nature of their dream.

Entrepreneurial characteristics and competences (such as initiative, leadership, creativity, perseverance, commitment, and imagination) cannot be mechanically replicated to construct an entrepreneur as if they were a robot; rather, the individual must be motivated to self-construct. To achieve this, emotional involvement is essential, enabling the student to generate such behaviours as the pursuit of dream fulfilment (Dolabela, 2003).

Since emotion is the driving force that stimulates and triggers intelligence and the creative process, the dreamer must define their dream at the very outset of the process. This dream should correspond to their aspirations for self-actualisation, and they must seek to realise it while managing the energy of the forces supporting this attainment. The strategy of entrepreneurial pedagogy is founded primarily on stimulating the dream. Subsequently, regarding the construction and realisation of this dream, educators should not interfere; instead, they should merely offer supporting elements to develop the competences and skills required to formulate and pursue it. Finally, guidance in delineating the ethical contours within which this entire activity will unfold is highly encouraged (Dolabela, 2003).

The dream generates the emotion that stimulates the desire for knowledge. If a specific content or process bears no relation to a student's dream, three hypotheses may explain this occurrence: denying meaning to an individual's dream (which constitutes an authoritarian, limiting, and non-educational act), denying meaning to knowledge itself, or a lack of connection between knowledge and its practical application (the latter being the most probable cause). Rejecting dreams is precisely the practice that entrepreneurial pedagogy intends to transform. Entrepreneurial knowledge can be categorised into four pillars: learning to know, learning to do, learning to be, and learning to live together (Dolabela, 2003).

As reality undergoes constant transformations throughout an individual's life, demanding a redefinition of the relationships between the dream, the pursuit of realisation, and supporting elements, this dynamic may require a learning process that restarts daily, inevitably leading to errors. These errors are viewed as sources of learning rather than as failures or grounds for capitulation. The connection between dreaming (in constant mutation) and striving to realise the dream (in constant evolution) constitutes the essence of the process. The relationship between “dreaming” and “realising the dream” is more critical than either element in isolation. When this relationship is proactive and dynamic, it generates both the necessity and the capacity for knowledge acquisition (Dolabela, 2003).

A dream transformed into actionable objectives is termed a vision. Every viable dream contains a vision, which can be shaped by various factors, including maturity, knowledge acquisition, beliefs, and values. Individuals with similar dreams develop distinct visions due to their personal biographies. In attempting to realise a dream, individuals expand their knowledge regarding the object of that dream and its surrounding context. Consequently, they test their skills and competences, becoming capable of identifying what they enjoy doing. When a dream is realised, the individual moves past transient emotions and begins to cultivate the capacity for perseverance in the face of errors and difficulties. Knowing their dream provides entrepreneurs with the impulse and energy to seek its fulfilment and commence their journey. Students must seek to understand what is required to realise their dream; this pursuit serves as the generative source of the impulse towards knowledge acquisition, which comprises four phases: learning to be, learning to do, learning to live together, and learning to know (Dolabela, 2003).

It is through individuality that an individual constructs their representation of the world. Individuality does not imply individualism or isolation, but rather the recognition of diversity—the capacity of each individual to construct distinct dreams, differentiated by their worldview and way of being. Although the entrepreneur conceives the dream independently, its realisation is impossible without the assistance of others and a supportive environment (Dolabela, 2003).

## Research Method

This study employs a descriptive research methodology, which is characterised by the observation, recording, analysis, description, and correlation of facts or phenomena without active manipulation. It seeks to ascertain with precision the frequency with which a phenomenon occurs and its relationship with other factors (Cervo and Bervian, 2002).

Descriptive research can assume various forms depending on the specific focus the researcher intends to establish for the study. Cervo and Bervian (2002) classify these forms as follows: exploratory studies, descriptive studies, survey research, and case studies. The present study

is classified under "descriptive studies." These studies describe the characteristics, properties, or existing relationships within the group or the specific reality in which the investigation was conducted.

To collect data for this study, the official website of the INEP (Instituto Nacional de Estudos e Pesquisas Educacionais) [National Institute for Educational Studies and Research] was utilised, as it provides a comprehensive registry of all recognised higher education courses in Brazil. Based on data retrieved from the INEP and the websites of the respective educational institutions, a spreadsheet was developed containing the following information:

- The city and state where the course is conducted
- The name of the institution
- The duration of the course
- The shift/period of instruction (e.g., day, evening)
- The number of student places available
- The curricular emphasis or specialisation
- The modality (e.g., conventional, distance learning)
- The academic calendar regime (e.g., semester, annual)
- The date on which the course commenced operations
- The number of modules/disciplines related to entrepreneurship
- The official quality score/rating
- The minimum required workload hours
- The workload hours allocated to core areas (Basic, Specific, Electives)

The primary focus of this research is to identify whether production engineering courses in Brazil incorporate modules related to the theme of entrepreneurship within their curriculum grids. To achieve this, the curricula of all 805 production engineering courses identified in Brazil were analysed.

## Results

The results indicate a strong concentration of programmes in the Southeast region, representing 61% of the total. The study also identified regional differences in the distribution of entrepreneurship-related subjects and highlighted the significant number of recently created programmes. Out of 805 programmes analysed, 443 had not yet graduated their first cohort, corresponding to fifty-five per cent, which demonstrates the relatively recent expansion of Industrial Engineering education in Brazil. This rapid growth reflects the increasing demand for professionals in this field. Of the 362 programs found, 150 Production Engineering programs do not have a Ministry of Education (MEC) score because the students from the first cohort have not yet graduated (based on data observed up to October 2025). Of the 212 programs that provided their curriculum (96 did not provide it, so their courses were not analyzed), it can be observed that 13% do not have any entrepreneurship-related courses (non-existent). Of the 87% that do cover the topic, 38% have a specific course named "Entrepreneurship" and 49% have entrepreneurship-related courses.

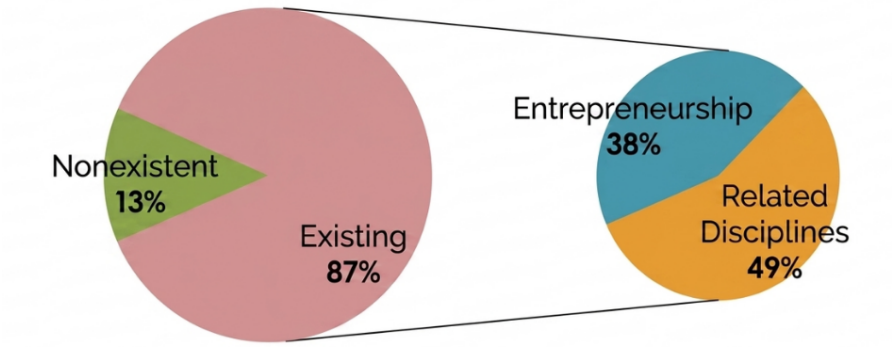
**Table 1**
*Distribution of Industrial Engineering Courses by Region in Brazil*

| Region | Percentage |
|---|---|
| North | 4% |
| Midwest | 5% |
| Northeast | 12% |
| South | 18% |
| Southeast | 61% |

*Source:* Field research (2025)

Out of the 212 courses that provided their curriculum grids (96 did not provide this information and their modules were consequently not analysed), it is observed that 13% do not offer any modules related to entrepreneurship (non-existent). Among the 87% that address the subject, 38% feature a specific module entitled "Entrepreneurship," whilst 49% offer modules related to entrepreneurship.

**Figure 1**
*Disciplines Related to the Theme of "Entrepreneurship"*




*Source:* Field research (2025)

Among the institutions offering the highest number of entrepreneurship-related modules, UNISINOS (Universidade do Vale do Rio dos Sinos) ranks first, offering 10 modules related to the theme, followed by UVA (Universidade Veiga de Almeida) with 7 modules. This information is detailed in Table 2.

**Table 2**
*Ranking of Institutions Offering the Highest Number of Entrepreneurship-Related Courses*

| Institution | Quantity | Institution | Quantity |
|---|---|---|---|
| UNISINOS | 10 | FUMEC | 5 |
| UVA | 7 | FTE | 5 |
| UEPA | 6 | UNIRADIAL-Curitiba | 5 |
| CEFET-RJ | 6 | UFERSA | 5 |
| UNESA | 6 | PUC-RS | 5 |
| UNIPLAC | 6 | IST-Florianópolis | 5 |
| UNIVATES | 6 | FAMEG | 5 |
| PUC-PR | 5 | UNISUL | 5 |
| UNIFOR | 5 | UNIRADIAL | 5 |
| UNIUBE | 5 | Others | 4, 3, 2 ou 1 |

*Source*: Field research (2025)

Regarding the average number of modules per region, it can be observed that the South and Northeast regions present the highest national averages of modules related to entrepreneurship, with 2.84 and 2.65 modules per course, respectively. Conversely, the Central-West region ranks lowest, offering an average of 2.13 modules per course.

The national average of entrepreneurship-related modules per production engineering course in Brazil stands at 2.44 modules per course. As previously demonstrated, the Central-West and Southeast regions fall below this benchmark. This disparity indicates that production engineering courses situated within the states and the Federal District of these regions must revise their entrepreneurship education to align with wider Brazilian developments.

Concurrently, this does not imply that courses located in the South, North, and Northeast regions should merely become complacent with their respective outcomes. Entrepreneurship is a highly contemporary subject that must be introduced to prospective undergraduate students of production engineering courses. This ensures they are equipped with a broader spectrum of activities to develop in their future professional careers, extending beyond the conventional employment options traditionally offered by the labour market.

## Conclusion

Production Engineering is a highly contemporary field of study in the country, as evidenced by the findings of this research, which indicate that 55.5% of the courses were established less than five years ago. Such a high growth rate results from the substantial market demand for professionals within this area of expertise.

Furthermore, it is possible to discern the extent to which higher education institutions incorporate the subject of entrepreneurship into their production engineering courses in Brazil, as well as which regions are more advanced in this regard. A significant insight derived from the research data highlights that the UFPR (Federal University of Parana) the focal object of this analysis offers a considerably traditional course compared to those of other Brazilian institutions. When contrasting the curriculum of this university with its national counterparts, the most notable deficiencies lie within modules covering administration, entrepreneurship, management, governance, innovation, leadership, and marketing. To elevate the standards of entrepreneurship education, it is imperative to evaluate the modules offered to students across the basic, specific, and elective components of the curriculum.

## Acknowledgements

The authors would like to thank the Specialization Program in Occupational Safety Engineering at the Federal University of Paraná (UFPR) for the institutional support and academic foundation provided.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors used the Gemini software during the preparation of this manuscript. The software was utilized solely for the purpose of language editing and proofreading to improve the clarity and readability of the text. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the final context of the published article.

## References

Cervo, A. R., & Bervian, P. A. (2002). *Metodologia científica* [Scientific methodology] (5th ed.). Prentice Hall.

Dolabela, F. (2003). *Pedagogia empreendedora* [Entrepreneurial pedagogy]. Editora de Cultura.

Filion, L. J. (1991a). *Visions et relations. Clefs du succès de l'entrepreneur* [Visions and relations. Keys to entrepreneur success]. Les Éditions de l'Entrepreneur.

Filion, L. J. (1991b). O planejamento do seu sistema de aprendizagem empresarial: Identifique uma visão e avalie o seu sistema de relações [Planning your business learning system: Identify a vision and evaluate your relationship system]. *Revista de Administração de Empresas, 31*(3), jul/set.

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira [INEP]. (2008). INEP. Ministério da Educação.

Pellin, C. M., & Santos, A. P. L. (2008). *O ensino do empreendedorismo no curso de engenharia de produção* [The teaching of entrepreneurship in the production engineering course]. In IV Encontro Mineiro de Engenharia de Produção. Ouro Preto, MG, Brasil.

**Contact emails:** adrianapls@ufpr.br
amadyta2004@gmail.com
brunolacerda99@gmail.com

# "It's a Private and Shameful Thing": Young People With Intellectual Disabilities Making Sense of Unwanted Sexual Messages on Social Media

Cansu Bolat Dinc, Marmara University, Turkiye
Fatma Nur Akcin, Marmara University, Turkiye

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

This study examines how young people with mild intellectual disability make sense of unwanted sexually explicit messages on social media and how they perceive related online sexual risks. Using a qualitative phenomenological design, semi-structured interviews were conducted with 30 young people aged 14–24 enrolled in a special education high school in Turkiye. The analysis explored participants' actual or anticipated responses to unwanted sexual images and messages, as well as the meanings they attached to privacy, shame, risk, and help-seeking. Findings show that participants described varied responses, including deleting or closing the image, blocking or unfollowing the sender, reporting or showing the content to family members, confronting the sender, remaining silent, or, in rare cases, responding in ways that could reproduce risk. These responses were shaped by emotional discomfort, family expectations, fear of blame, moral judgement, bodily privacy, and shame. The study suggests that online sexual risk for young people with mild intellectual disability should be understood as an educational, relational, and safeguarding concern. It highlights the need for accessible, safeguarding-informed online safety education using concrete scenarios, simple language, visual supports, and trusted-adult pathways.

*Keywords:* intellectual disability, social media, sexual messages, online safety, privacy

iafor
The International Academic Forum
www.iafor.org

## Introduction

Intellectual disability is generally understood as involving limitations in intellectual functioning and adaptive behaviour, including daily independence, social participation, and responsibility (American Association on Intellectual and Developmental Disabilities [AAIDD], 2010). However, disability cannot be understood only through individual functioning; it is also shaped by social conditions, cultural expectations, and access to support (Shakespeare, 2016). Within this context, privacy is a fundamental right connected to dignity, autonomy, and self-protection (Couwenhoven, 2007). For young people with intellectual disabilities, recognising boundaries between private and public information, bodies, relationships, and online interactions becomes particularly relevant when navigating social media and making sense of sexually explicit messages (Brown et al., 2024; eSafety Commissioner, 2020).

Social relationships are central to young people's identity, belonging, and emotional well-being (Ryan & Deci, 2000). Belonging is also shaped by experiences of acceptance, connection, and participation within social environments (Davies et al., 2025). For young people with intellectual disabilities, forming and maintaining relationships may be influenced by cognitive, communicative, social, and environmental barriers (Bishop et al., 2024). As social media has become embedded in young people's everyday lives, it can function as an important space for friendship, self-expression, belonging, and relationship-building (Anderson & Jiang, 2018). Digital platforms may offer opportunities for visibility and connection, yet they may also create complex situations involving sexuality, privacy, consent, and online sexual exploitation (ECPAT International, 2025; Henry et al., 2020). These opportunities and risks highlight the need for online safety approaches that not only protect young people from harm but also support the autonomy, participation, and digital citizenship of young people with intellectual disabilities (Bolat Dinc & Done, 2025). Against this background, previous studies have examined internet use, social media, online safety, and support among young people with intellectual disabilities (Borgström, 2023; Borgström et al., 2019; Caton & Landman, 2022). However, less is known about how young people with mild intellectual disability themselves interpret sexually explicit messages on social media, particularly through everyday concepts such as privacy, the private, shame, risk, and appropriate social behaviour (Borgström et al., 2019; Hart et al., 2025).

Online sexual violence may involve a range of non-contact behaviours, including unwanted sexually explicit messages and images, as well as related forms such as image-based sexual abuse, sextortion, and grooming (National Society for the Prevention of Cruelty to Children [NSPCC], n.d.). These risks require particular attention in relation to children and young people with disabilities, not because vulnerability is inherent to disability, but because social, communicative, educational, and support-related inequalities may shape exposure to harm and access to protection (United Nations Children's Fund [UNICEF], 2023). Research shows that unwanted sexual messages and images can lead to strong emotional responses, including shame, fear, anger, guilt, and insecurity, and may be associated with depression, anxiety, low self-esteem, self-blame, social withdrawal, and suicidal ideation (Mandau, 2020; Marcotte et al., 2021; Page et al., 2025). Such findings suggest that non-contact digital sexual abuse cannot be treated as less serious simply because it does not involve physical contact (Page et al., 2025).

This study examines how young people with mild intellectual disability make sense of sexually explicit messages on social media, how they describe their possible responses to such messages, and how concepts such as privacy, private, shame, and risk shape their

interpretations. In doing so, the study contributes to research on online sexual risks, intellectual disability, and the development of accessible online safety tools. Using a qualitative phenomenological approach, the study explored the experiences and narratives of 30 participants aged 14–24, including 15 girls and 15 boys. The research questions are as follows:

i. How do young people diagnosed with mild intellectual disability describe the possibility of encountering unwanted sexually explicit messages on social media and their reactions to such messages?
ii. How are the risk perceptions of these young people regarding sexually explicit messages they may encounter on social media shaped?

Overall, the study seeks to move beyond a risk-focused understanding of online sexual harm by examining how young people with mild intellectual disability themselves interpret, describe, and negotiate sexually explicit messages in relation to privacy, shame, and social boundaries. By foregrounding their own perspectives, the study aims to inform more accessible, rights-based, and context-sensitive approaches to online safety education and support.

## Research Methods

This study adopted a qualitative phenomenological approach to examine how young people aged 14–24 with mild intellectual disability make sense of sexually explicit messages on social media, how they describe their possible responses to such messages, and how they perceive related risks. The study was grounded in an interpretivist epistemological position, which understands qualitative research as a way of exploring participants' experiences, meanings, and subjective realities within their social, relational, and cultural contexts (Denzin & Lincoln, 2018). Phenomenology was used because the study was concerned not only with what participants said they might do when encountering sexually explicit messages, but also with how they interpreted such experiences through concepts such as privacy, private, shame, risk, and social boundaries (Creswell & Poth, 2018).

The study included 30 young people with mild intellectual disability who were enrolled in a special education high school in Turkiye. The sample comprised 15 girls and 15 boys aged between 14 and 24. Although some participants were over the age of 18, they were included because they were still receiving special education at high-school level and continued to share similar school-based, peer-related, and adolescence-linked experiences with younger students in the same educational setting. The wider age range also reflected delayed diagnosis, later school entry, extended education pathways, and disruptions to schooling associated with the COVID-19 pandemic. For this reason, the term young people are used throughout the study rather than children alone. Demographic information is presented in Table 1.

**Table 1**
*Demographic Information Regarding Participants*

| Participant ID | Gender | Grade level | Age (years) | Diagnosis |
|---|---|---|---|---|
| P01 | Male | 9th grade | 15 | MID |
| P02 | Male | 9th grade | 15 | MID |
| P03 | Male | 9th grade | 18 | MID |
| P04 | Male | 10th grade | 18 | MID |
| P05 | Male | 10th grade | 17 | MID |
| P06 | Male | 10th grade | 17 | MID |
| P07 | Male | 10th grade | 18 | MID |
| P08 | Male | 10th grade | 18 | MID |
| P09 | Male | 10th grade | 19 | MID |
| P10 | Male | 11th grade | 18 | MID |
| P11 | Male | 11th grade | 18 | MID |
| P12 | Male | 11th grade | 19 | MID |
| P13 | Male | 12th grade | 19 | MID |
| P14 | Male | 12th grade | 19 | MID |
| P15 | Male | 12th grade | 21 | MID |
| P16 | Female | 9th grade | 14 | MID |
| P17 | Female | 9th grade | 16 | MID |
| P18 | Female | 10th grade | 16 | MID |
| P19 | Female | 10th grade | 16 | MID |
| P20 | Female | 10th grade | 16 | MID |
| P21 | Female | 10th grade | 17 | MID |
| P22 | Female | 10th grade | 17 | MID |
| P23 | Female | 10th grade | 17 | MID |
| P24 | Female | 10th grade | 19 | MID |
| P25 | Female | 10th grade | 20 | MID |
| P26 | Female | 11th grade | 18 | MID |
| P27 | Female | 11th grade | 19 | MID |
| P28 | Female | 11th grade | 21 | MID |
| P29 | Female | 11th grade | 22 | MID |
| P30 | Female | 11th grade | 24 | MID |

*Note.* MID (Mild Intellectual Disability)

Criterion sampling, a purposeful sampling strategy, was used to select participants who met predefined criteria relevant to the study aim (Creswell & Poth, 2018). The inclusion criteria were diagnosis of mild intellectual disability, attendance at a special education institution at high-school level, experience of or potential exposure to social media use, and written consent from the participant and/or their parent or legal guardian. The study centred young people's own accounts because they were regarded as experts in their own lived experiences, rather than as clinical or technical experts on online sexual harm. All interviews were conducted individually and face-to-face in a quiet room within the school setting. The semi-structured interview format allowed participants to express their views in their own words, while also enabling the researcher to clarify questions, provide reflection time, and ask follow-up questions when necessary (Gay et al., 2011; Patton, 2002). An interview guide was developed through a literature review, reviewed by six experts, revised according to their feedback, and finalised after pilot interviews.

Qualitative data analysis involved pre-analysis preparation and thematic interpretation. All interviews were audio-recorded, securely backed up, and transcribed verbatim; significant verbal and behavioural responses were also noted. Transcripts were read repeatedly to develop familiarity with the data, and initial coding focused on how participants described sexually explicit messages, identified risk, described possible responses, and used concepts such as privacy, private, shame, body, speech, behaviour, and environment. The analysis was both conceptually informed and data-driven: the initial framework was shaped by the study's focus on privacy, social boundaries, and online sexual risk, while sub-themes were developed through close engagement with participants' own narratives (Miles & Huberman, 1994).

To ensure consistency with the interpretivist and phenomenological orientation of the study, coding was treated as a reflexive and interpretive process rather than as a procedure requiring statistical inter-rater reliability. Discussions with a field expert were used to refine interpretations, question assumptions, and strengthen analytic depth. Trustworthiness was supported through verbatim transcription, repeated reading of the data, reflexive memo-writing, expert consultation, transparent documentation of analytic decisions, and the use of participant quotations to ground the findings.

### Ethical Considerations

Ethical approval was obtained before the study, and written consent was received from participants, parents/guardians, and institutional administrators. Participation was voluntary, and participants were informed of their right to withdraw at any time. One male participant withdrew during the interview; the interview was terminated immediately, and his data were excluded from the analysis. Interviews were successfully completed with 30 participants. Code names were used to protect identities, and all data were securely stored and used only for academic purposes. Any disclosures of abuse were referred to the institution's guidance counsellors without sharing audio recordings.

## Results

This section presents findings related to how young people with mild intellectual disability described actual or anticipated encounters with unwanted sexually explicit messages on social media and their responses to such encounters (RQ1). It then examines how these young people constructed and negotiated perceptions of online sexual risk, particularly through ideas of shame, privacy, discomfort, moral judgement, and possible harm (RQ2).

### Encountering Unwanted Sexual Messages on Social Media: Response Repertoires (RQ1)

When participants' statements regarding how they reacted to inappropriate photos sent to them via social media accounts were examined, it was observed that the reactions clustered on multiple levels. While some young people focused on ending the situation with direct actions against the photo and the account (deleting, closing, blocking, reporting), some mentioned a tendency to share the situation with family members or other adults, or to react directly to the sender. A smaller number of participants emphasized their emotional reactions or expressed a tendency not to tell anyone. These findings are summarized in Table 2 in the form of codes and frequencies.

**Table 2**
*Participants' Responses to Unwanted Sexually Explicit Images Received via Social Media*

| Theme | Subtheme | f |
|---|---|---|
| **Reactions given via social media account** | | **35** |
| | I would delete the photo and close it. I wouldn't look at the photos. | 12 |
| | I would block them. | 12 |
| | I would mark it as inappropriate in the app and report it. | 7 |
| | I would unfollow them. I would delete their number. | 3 |
| | I would send them a photo of my own bottom. | 1 |
| **Reactions involving telling others about the situation** | | **16** |
| | I would tell my mum and dad. | 7 |
| | I would report it to the police. | 6 |
| | I would tell my teacher. | 1 |
| | I would tell my girlfriends. | 1 |
| | I would show the photos to my older sister. | 1 |
| **Reactions towards the person who sent the photos** | | **13** |
| | I would slap them. I would beat them up. | 5 |
| | I would say, "It's shameful, don't send it." | 3 |
| | I would swear at them. | 2 |
| | I would never speak to them again. | 2 |
| | I would shout at them. | 1 |
| **Emotional responses** | | **4** |
| | I would get angry. | 2 |
| | I would feel sad / I would cry. | 2 |
| **Other** | | **2** |
| | I wouldn't tell anyone. My family would be very angry. | 2 |
| **Total responses** | | **70** |

One of the most described responses was to delete the image, close the application, or avoid looking at the content. Twelve participants (40%) referred to this type of immediate distancing. For these participants, deleting or closing the image appeared to function not only as a digital safety action but also as a way of removing themselves emotionally and socially from an uncomfortable situation.

> I get angry, I close it immediately. (P24)

> What would I do, I'd delete them. … Umm, I would remove them. … I wouldn't tell my mother anything. I wouldn't tell my family or anyone else. … If I told them, my parents would kill me, that's why. (P23)

> I'd call them a pervert. It's shameful, disgraceful. I would close that picture too. You shouldn't open this when you're around other people. (P14)

These accounts suggest that deleting, closing, or avoiding the image was not only a way to escape the content, but also a strategy to prevent others from seeing it and to avoid possible blame, anger, or punishment from family members. Similarly, blocking, mentioned by 12 participants (40%), functioned as a form of boundary-making. It was often linked with unfollowing, reporting, closing the sender's profile, or expressing anger, reflecting participants' attempts to stop further contact, regain control, and mark the sender's behaviour as unacceptable.

> I would block them straight away. … Umm, and I would also unfollow that person, and even if I'm not following them, I would completely close their page. Like, umm, I wouldn't swear but I would send messages like 'I hope your mother or your sister sees the same thing.' … I'd want them to go through the same thing, to feel the same way. (P30)

> I would block them, delete it without looking. … I would report it. There's nothing else to do. If it's someone I know, they'd get beaten up. (P04)

Blocking was not understood only as a platform function; for some participants, it was also linked to punishment, confrontation, complaint, or holding the sender accountable for a social or moral offence. Telling family members was another response, mentioned by seven participants (23.3%), and was sometimes connected with making an official complaint, including going to the police.

> To my mum or dad. … I would tell them. … And then, I could go to file a complaint. … Umm, to the police officers. (P20)

> I would show it to my mum. … I wouldn't show it to anyone else. (P21)

Family members were positioned as potential sources of protection, yet disclosure was not always straightforward. Some participants feared blame, punishment, or negative family reactions, suggesting that help-seeking could be shaped by shame and fear as well as safety. In addition, five participants (16.6%) described physical retaliation, such as slapping or beating the sender, indicating that some understood online sexual transgression through face-to-face norms of disrespect, anger, and accountability.

> I would slap them once. (P16)

> I'd get angry. … I'd beat them up. (P12)

For these participants, the digital violation was not treated as separate from offline social relations. Instead, the sender's behaviour was imagined as something that could require direct confrontation or punishment in physical space. This finding suggests that online sexual harm may be understood by some young people through familiar social scripts of disrespect, anger, and face-to-face conflict.

A smaller number of participants described responses that could reproduce or intensify risk. One participant stated that he would send a photo of his own private parts in return:

> I would send one to them too. … A photo. … Umm of my private parts. … Umm my bum. (P07)

This response suggests that not all participants interpreted the safest or most protective course of action in the same way. In some cases, the idea of responding *in the same way* may reflect limited understanding of the consequences of reciprocal image-sharing, especially in relation to privacy, consent, and digital permanence.

One of the most detailed narratives came from a female participant who described both receiving sexually explicit images and previously sending an image of her own private parts.

Her account shows how family responses, shame, privacy, and retrospective understanding of boundaries became intertwined:

> I show it straight to my sister. Because I've experienced it a few times. … I just said, sis come here, I'm going to show you something. I showed her the photo. That's it. Then my sister said, who is this boy. She said, tell me right now. She said, I'm going to file a complaint. That's it. … I mean. She said, just send it back to him. I did it once. But I did it by mistake. … My sisters got angry with me. They said, why did you do something like that. She said that's a very private area. … I said yes, I know that too. She didn't talk to me for a month. That's why. (P29)

This account brings together exposure to unwanted sexual images, help-seeking through an older sister, the possibility of complaint, and the participant's later understanding of her own action as a boundary violation. The phrase *a very private area* illustrates how family discourses of privacy shaped her interpretation, while the sister's anger and withdrawal suggest that shame and family judgement may complicate help-seeking. Overall, participants' responses were multifaceted and shaped by emotional reactions, family expectations, perceived social consequences, shame, and bodily privacy. These findings indicate that unwanted sexually explicit messages were interpreted not only as digital incidents, but also through family, social, emotional, and moral contexts.

**Making Sense of Online Sexual Risks: Dimensions of Risk Perception (RQ2)**

Participants' responses regarding the reasons for their reactions to inappropriate photos sent via social media accounts offer clues about the dimensions around which risk perception is shaped. These responses are grouped under the headings of their own value judgments (describing them as wrong, bad, inappropriate, shameful, indecent, and perverted), emotional reactions (discomfort, dislike, not wanting to see them), reasons related to the person they reported the harassment to, and concerns about digital security. Frequency distributions are shown in Table 3.

**Table 3**
*Reasons Shaping Participants' Responses to Unwanted Sexually Explicit Images on Social Media*

| Theme | Subtheme | f |
|---|---|---|
| **Reasons related to own moral judgments** | | **18** |
| | Wrong / bad / inappropriate behaviour | 8 |
| | Shameful / disrespectful things | 7 |
| | They might be a pervert. | 3 |
| **Emotional reasons** | | **13** |
| | I would feel uncomfortable / I wouldn't like it / I wouldn't want to see it. | 7 |
| | I don't want them to send such photos. | 6 |
| **Reasons related to the person they disclose the abuse to** | | **3** |
| | I would tell my mother that she is the first to know. | 2 |
| | I would tell my older sister because she trusts me. | 1 |
| **Other** | | **2** |
| | My social media account could be stolen. A virus could infect it. | 2 |
| **Total responses** | | **36** |

A few participants framed sexually explicit images as wrong, bad, disrespectful, or inappropriate. Eight participants (26.6%) explained their reactions through this type of moral and normative judgement.

> These are bad things, teacher. … They're doing something wrong, something disrespectful. (P28)

> Because they're very bad photos. … Because they're disgusting. (P06)

> They send bad things, things I don't want. … I don't like things like that. I don't look at anything like that. (P04)

In these accounts, participants framed the images not merely as unwanted, but as morally unacceptable and socially disrespectful. Terms such as bad, wrong, disgusting, and disrespectful functioned as everyday moral categories that justified actions such as deleting, blocking, or reporting. Shame was also central: seven participants (23.3%) described sexually explicit images as shameful or indecent, suggesting that risk was understood not only through danger or security, but also through social norms about what should or should not be seen, shown, or shared.

> Because they did something very shameful. (P27)

> Because they send that kind of stuff, teacher, they send shameful things, that's why. I would block them. (P13)

> Because these are perversion, lack of manners, disrespect. After that, there can be rape, violence, anything. (P11)

Shame, indecency, and disrespect were sometimes linked to more serious risks such as abuse, aggression, and physical danger. In this sense, online sexual content was not always understood as separate from offline harm. Emotional discomfort also shaped risk perception; seven participants (23.3%) explained their reactions through feelings such as discomfort, anger, dislike, or not wanting to see the image. Here, risk was experienced not as an abstract issue, but as an immediate emotional and bodily response.

> So, I would block them because I don't want to see it. (P26)

> I wouldn't want them to send photos like that. … I wouldn't talk to them again. (P08)

These statements suggest that unwanted sexually explicit images disrupted participants' sense of emotional safety. The images were not only judged as wrong; they were also experienced as unpleasant, disturbing, and unwanted. Blocking or cutting off communication therefore became a way of protecting emotional boundaries as well as digital boundaries.

One participant described this emotional reaction in more detail:

> No, once there were people who sent that to me once. … It was, you know, his thing as a boy. That's it. I mean, a boy had sent me a photo or something. … And I got really angry. … I closed it immediately. … That's it. I mean, that boy, like, I have my account

> and he writes stupid stuff. 'Where are you, how are you?' And stuff like, 'Did you like it?' and so on. … Then I closed it, I got angry. … That's it. (P24)

This narrative shows that sexually explicit images were experienced not as isolated incidents, but as part of wider unwanted interactions involving repeated messages, intrusive questions, and pressure to respond. The participant's anger and decision to close the interaction can be understood as an attempt to regain control. Across the accounts, risk perception was shaped by moral language, shame, emotional discomfort, and concerns about escalation. Shame helped participants recognise such messages as inappropriate, but it could also make disclosure difficult when combined with fear of family anger, judgement, or blame.

Overall, young people with mild intellectual disability did not describe unwanted sexually explicit messages only as technical problems to be solved through platform functions. They interpreted them through everyday concepts such as bad, shameful, private, disgusting, disrespectful, and dangerous. These meanings shaped how they responded, whether they blocked or deleted the content, whether they told someone, and whether they imagined the incident as something that could remain online or move into face-to-face social life. This indicates that online safety support for young people with mild intellectual disability needs to address not only technical actions such as blocking and reporting, but also the emotional, moral, relational, and shame-related meanings through which young people make sense of online sexual risk.

## Discussion

The literature includes a limited number of studies directly examining how young people with intellectual disability respond to unwanted sexually explicit images or messages on social media. For this reason, the findings of this study can be interpreted in relation to broader discussions on sexual abuse, bodily privacy, digital risk, safeguarding, and online safety education. Showing or exposing someone to images of genitals or private body parts, taking or sharing unauthorised images, peeping, and using sexually suggestive language are recognised as non-contact forms of sexual abuse (NSPCC Learning, 2026; UNICEF, 2023). Digital environments may intensify these risks because unwanted sexual content can be sent quickly, privately, repeatedly, and in ways that may be difficult for young people and adults to notice or control (Livingstone & Smith, 2014; Weeks, 2017). In educational terms, this indicates the need to understand online sexual harm not only as a technological issue, but also as a safeguarding and learning issue that requires accessible teaching about privacy, consent, help-seeking, and safe responses (Brown et al., 2024; eSafety Commissioner, 2020).

Previous research suggests that individuals with intellectual disabilities often desire close friendships, intimacy, and social participation, but may experience barriers in distinguishing appropriate and inappropriate interpersonal behaviours (Swango-Wilson, 2011). International studies also indicate that young people with intellectual disability may experience difficulties recognising boundary violations in online contexts, which may increase the risk of abuse (Normand & Sallafranque-St-Louis, 2016; Stöckl et al., 2014). The findings of this study both support and complicate this literature. Many participants described rejecting responses, such as deleting images, blocking senders, reporting accounts, or telling family members. However, these responses were not always embedded in a clearly protective framework. Some accounts suggested uncertainty about how to respond safely, particularly when participants imagined retaliating, responding in kind, or keeping the experience secret. This suggests that online safety education may need to move beyond general warnings and provide concrete, scenario-

based learning opportunities that help young people recognise boundary violations and practise safe responses (Department for Education, 2023; eSafety Commissioner, 2022).

Two participants' accounts were particularly important in showing this complexity. One female participant (P29) described receiving unwanted sexual images and, in another situation, sending an image of her own private parts. Similarly, one male participant (P07) stated that he would send a photo of his private parts in response to receiving such content. These examples suggest that some young people may not always clearly distinguish between rejecting an abusive situation, retaliating, and reproducing risk. From an educational perspective, this finding highlights the importance of teaching privacy, consent, and digital permanence through simple, concrete, and repeated examples. Educational materials on online safety and relationships could use accessible scenarios to explore questions such as what counts as a private body part, why sending images of private areas can create risk, how to say no, and whom to tell when something uncomfortable happens online.

The findings also show that emotional reactions to unwanted sexual content do not necessarily lead to help-seeking. For example, one female participant (P24) described feeling intense anger and immediately closing the image but did not indicate that she would tell a trusted adult. In this study, participants often used every day moral and emotional terms such as bad, shameful, disgusting, disturbing, and disrespectful. These terms appeared to help them recognise the content as wrong but did not always translate into clear protective action. Online safety education may therefore be more effective when it connects formal safeguarding language with the everyday words young people use to describe discomfort, shame, fear, and risk. This reflects a wider issue in the literature, which suggests that limited knowledge of bodily privacy and sexual boundaries may make it difficult for individuals with intellectual disabilities to recognise, name, and report abuse (Noland, 2010).

Shame emerged as a particularly important theme. For some participants, shame helped define sexually explicit images as inappropriate and unacceptable. At the same time, shame could also make disclosure more difficult, especially when participants feared family anger, blame, or punishment. This dual role of shame is important for educational practice. Teaching young people to recognise private body parts and inappropriate sexual messages is necessary, but it may not be sufficient if they also learn that such experiences are too shameful to discuss. This suggests that privacy and safeguarding education need to include explicit reassurance that receiving unwanted sexual content is not the young person's fault, and that asking for help is an appropriate and protective response (Brown et al., 2024; Department for Education, 2025).

The role of trusted adults was also central. Some participants positioned mothers, fathers, siblings, or the police as possible sources of protection. However, others hesitated to disclose because they feared negative reactions from family members. This indicates that online safety education cannot focus only on young people's individual knowledge or behaviour. It also needs to consider the wider support environment around them, including families, teachers, school counsellors, and safeguarding staff. Prior research similarly emphasises the importance of trusted adults and supportive communication in helping young people with intellectual disabilities understand online risks and seek help (Holmes & O'Loughlin, 2014; Stone et al., 2017). In this sense, schools may have an important role in creating safe, non-judgemental spaces where young people can talk about uncomfortable online experiences without fear of blame.

These findings have implications for digital citizenship, relationship and sexuality education, and safeguarding education for young people with intellectual disability. Topics such as bodily privacy, consent, saying no, blocking and reporting, digital permanence, and help-seeking may need to be taught together through accessible, scenario-based activities. Visual stories, role-play, simple decision trees, and repeated practice may help young people recognise boundary violations and practise safe responses while supporting autonomy, dignity, and participation (Brown et al., 2024; Department for Education, 2023; eSafety Commissioner, 2022).

The findings also suggest the value of involving young people with intellectual disability in the development of online safety and relationship education materials. Their accounts show that they do not only need protection from risk; they also actively interpret, judge, resist, and make sense of online sexual content using their own moral, emotional, and relational frameworks. Educational resources may therefore become more relevant and accessible when they are informed by young people's own language, examples, concerns, and support preferences. Co-designed materials with young people, educators, families, and safeguarding professionals could help ensure that online safety education is not only technically accurate, but also emotionally meaningful, developmentally accessible, and sensitive to shame, privacy, and family dynamics (Brown et al., 2024; Marsden et al., 2022; Wellington et al., 2025).

In summary, this study extends existing discussions on online sexual risk among young people with intellectual disability by showing that responses to unwanted sexually explicit messages are shaped by more than awareness of danger. Participants' accounts were also shaped by shame, family expectations, emotional discomfort, moral judgement, bodily privacy, and uncertainty about appropriate action. For education, this points to the need for accessible, rights-based, and context-sensitive online safety education that addresses both knowledge and emotion, both individual decision-making and trusted support, and both digital risks and the social meanings through which young people understand them (Brown et al., 2024; eSafety Commissioner, 2020; UNICEF, 2024).

## Conclusion

The findings of this study indicate that young people with mild intellectual disabilities make sense of unwanted sexually explicit messages on social media through a complex interplay of technical, emotional, moral, relational, and privacy-related meanings. Their responses, including deleting, blocking, reporting, telling family members, confronting the sender, or remaining silent, were not merely practical reactions to online content; they also reflected broader concerns about shame, bodily privacy, family judgement, social exposure, and possible harm. The analysis suggests that these young people often recognised such messages as wrong, shameful, disturbing, or dangerous, yet help-seeking could be shaped by fear of blame, punishment, or being misunderstood. By foregrounding their own accounts, the study highlights the importance of understanding online sexual risk not only through digital safety functions, but also as an experience embedded in relationships, emotions, social norms, and accessible meanings of privacy and protection. These findings point to the need for online safety education that combines digital skills with accessible teaching about privacy, consent, emotional safety, and trusted help-seeking.

## Limitations and Future Research

This study has several limitations that may guide future research. The findings are based on a purposive sample of 30 young people with mild intellectual disability from a special education

high-school context in Turkiye; therefore, the results cannot be generalised to all young people with intellectual disabilities or to different cultural and educational settings. The study also relied on participants' self-reported accounts of actual or anticipated responses to unwanted sexually explicit messages, and the sensitivity of the topic may have shaped what participants felt able or willing to share. In addition, the study centred young people's own perspectives and did not include the views of parents, teachers, counsellors, or digital safety professionals, whose perspectives could provide further insight into support practices and safeguarding responses. Future research could include more diverse participant groups, different levels of intellectual disability, cross-cultural comparisons, and longitudinal or participatory designs. Co-designed studies with young people with intellectual disability may also help develop more accessible online safety tools, scenario-based educational materials, and support pathways that respond to the emotional, relational, and shame-related dimensions of online sexual risk.

## Acknowledgements

We sincerely thank the young people who participated in this study and shared their perspectives. We are also grateful to their parents and legal guardians for providing permission and supporting the research process. We further acknowledge the support of the Ministry of National Education of the Republic of Turkiye.

## Author's Note

This paper is based on research conducted as part of Cansu Bolat Dinc's master's thesis at Marmara University. She is currently a doctoral researcher at the University of Plymouth, United Kingdom.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that ChatGPT (GPT-5.6 Thinking, OpenAI), an AI-assisted writing tool, was used in proofreading, refining the language, and improving the organisation of the manuscript. Its use was limited to correcting grammatical and spelling errors, reorganising the text, and rephrasing statements for accuracy and clarity. The authors further declare that, apart from ChatGPT, no other AI or AI-assisted technologies were used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are the authors' original work and were derived from the careful and systematic conduct of the research.

## References

American Association on Intellectual and Developmental Disabilities. (2010). *Intellectual disability: Definition, classification, and systems of supports* (11th ed.).

Anderson, M., & Jiang, J. (2018, May 31). *Teens, social media & technology 2018*. Pew Research Center. https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/

Bishop, G. M., Llewellyn, G., Kavanagh, A. M., Badland, H., Bailie, J., Stancliffe, R., Emerson, E., Fortune, N., & Aitken, Z. (2024). Disability-related inequalities in the prevalence of loneliness across the lifespan: Trends from Australia, 2003 to 2020. *BMC Public Health, 24*, Article 621. https://doi.org/10.1186/s12889-024-17936-w

Bolat Dinc, C., & Done, E. J. (2025). From detection to protection: Online risks, privacy and security in the lives of disabilities. *International Journal for Digital Society, 16*(1), 2023–2028. https://doi.org/10.20533/ijds.2040.2570.2025.0252

Borgström, Å. (2023). Tensions between risk, coping and support: Young people with intellectual disability in Sweden and Internet-related support. *Disability & Society, 38*(3), 460–482. https://doi.org/10.1080/09687599.2021.1946675

Borgström, Å., Daneback, K., & Molin, M. (2019). Young people with intellectual disabilities and social media: A literature review and thematic analysis. *Scandinavian Journal of Disability Research, 21*(1), 129–140. https://doi.org/10.16993/sjdr.549

Brown, M., Linden, M., Marsh, L., Truesdale, M., Sheerin, F., & McCormick, F. (2024). Learning for life, friendships and relationships from the perspective of children and young people with intellectual disabilities: Findings from a UK wide qualitative study. *BMC Public Health, 24*, Article 2491. https://doi.org/10.1186/s12889-024-19972-y

Caton, S., & Landman, R. (2022). Internet safety, online radicalisation and young people with learning disabilities. *British Journal of Learning Disabilities, 50*(1), 88–97. https://doi.org/10.1111/bld.12372

Couwenhoven, T. (2007). *Teaching children with Down syndrome about their bodies, boundaries, and sexuality*. Woodbine House.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Davies, C. A., Cordier, R., Graham, P., Littlefair, D., Speyer, R., & Melo, D. (2025). Interventions to improve connectedness, belonging, and engagement in secondary schools: A systematic review and meta-analysis. *Education Sciences, 15*(5), Article 582. https://doi.org/10.3390/educsci15050582

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2018). *The SAGE handbook of qualitative research* (5th ed.). SAGE Publications.

Department for Education. (2023). *Teaching online safety in schools*. GOV.UK. https://www.gov.uk/government/publications/teaching-online-safety-in-schools/teaching-online-safety-in-schools

Department for Education. (2025). *Relationships education, relationships and sex education (RSE) and health education*. GOV.UK. https://www.gov.uk/government/publications/relationships-education-relationships-and-sex-education-rse-and-health-education

ECPAT International. (2025). *Terminology guidelines for the protection of children from sexual exploitation and sexual abuse*. https://ecpat.org/wp-content/uploads/2025/04/Second%20Edition-Terminology-Guidelines-final.pdf

eSafety Commissioner. (2020). *Online safety for young people with intellectual disability*. Australian Government. https://www.esafety.gov.au/research/online-safety-for-young-people-with-intellectual-disability

eSafety Commissioner. (2022). *Be an eSafe kid: Online boundaries—It's OK to say "no"*. Australian Government. https://www.esafety.gov.au/sites/default/files/2022-06/Online%20boundaries%20-%20Scenarios.pdf

Gay, L. R., Mills, G. E., & Airasian, P. W. (2011). *Educational research: Competencies for analysis and applications* (10th ed.). Pearson.

Hart, A., Carlström, C., Löfgren, C., & Elmerstig, E. (2025). Sexuality and intimacy among young adults with intellectual disabilities in Sweden: Support networks. *Journal of Social Work*. Advance online publication. https://doi.org/10.1177/14680173251383374

Henry, N., McGlynn, C., Flynn, A., Johnson, K., Powell, A., & Scott, A. J. (2020). *Image-based sexual abuse: A study on the causes and consequences of non-consensual sexual image sharing*. Routledge.

Holmes, K. M., & O'Loughlin, N. (2014). The experiences of people with learning disabilities on social networking sites. *British Journal of Learning Disabilities, 42*(1), 1–5. https://doi.org/10.1111/bld.12001

Livingstone, S., & Smith, P. K. (2014). Annual research review: Harms experienced by child users of online and mobile technologies: The nature, prevalence and management of sexual and aggressive risks in the digital age. *Journal of Child Psychology and Psychiatry, 55*(6), 635–654. https://doi.org/10.1111/jcpp.12197

Mandau, M. B. H. (2020). "Directly in your face": A qualitative study on the sending and receiving of unsolicited "dick pics" among young adults. *Sexuality & Culture, 24*, 72–93. https://doi.org/10.1007/s12119-019-09626-2

Marcotte, A. S., Gesselman, A. N., Fisher, H. E., & Garcia, J. R. (2021). Women's and men's reactions to receiving unsolicited genital images from men. *The Journal of Sex Research, 58*(4), 512–521. https://doi.org/10.1080/00224499.2020.1779171

Marsden, L., Moody, L., Nguyen, B., Tatam, L., Welland, L., & Third, A. (2022). *Reimagining online safety education through the eyes of young people: Co-design workshops with young people to inform digital learning experiences*. Western Sydney University. https://doi.org/10.26183/3bz3-r451

Miles, M. B., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). SAGE Publications.

National Society for the Prevention of Cruelty to Children. (n.d.). *Online sexual exploitation and abuse of children*. https://www.nspcc.org.uk/what-is-child-abuse/types-of-abuse/child-sexual-exploitation/

Noland, C. M. (2010). *Sex talk: The role of communication in intimate relationships*. Praeger.

Normand, C. L., & Sallafranque-St-Louis, F. (2016). Cybervictimization of young people with an intellectual or developmental disability: Risks specific to sexual solicitation. *Journal of Applied Research in Intellectual Disabilities, 29*(2), 99–110. https://doi.org/10.1111/jar.12163

NSPCC Learning. (2026). *Protecting children from sexual abuse*. National Society for the Prevention of Cruelty to Children. https://learning.nspcc.org.uk/child-abuse-and-neglect/child-sexual-abuse

Page, S., McFeeters, A., Mat Rani, N. H., Slater, C., Fry, D., Stevenson, J., & Quayle, E. (2025). Psychological and emotional harms of online child sexual exploitation and abuse in children under the age of 18: A systematic review of the evidence. *Child Abuse Review, 34*(4), Article e70048. https://doi.org/10.1002/car.70048

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). SAGE Publications.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Shakespeare, T. (2016). *Disability: The basics*. Routledge.

Stöckl, H., March, L., Pallitto, C., & Garcia-Moreno, C. (2014). Intimate partner violence among adolescents and young women: Prevalence and associated factors in nine countries: A cross-sectional study. *BMC Public Health, 14*, Article 751. https://doi.org/10.1186/1471-2458-14-751

Stone, N., Ingham, R., McGinn, L., & Bengry-Howell, A. (2017). Talking relationships, babies and bodies with young children: The experiences of parents in England. *Sex Education, 17*(5), 588–603. https://doi.org/10.1080/14681811.2017.1332988

Swango-Wilson, A. (2011). Meaningful sex education programs for individuals with intellectual/developmental disabilities. *Sexuality and Disability, 29*(2), 113–118. https://doi.org/10.1007/s11195-010-9168-2

United Nations Children's Fund. (2023). *International classification of violence against children*. https://data.unicef.org/resources/international-classification-of-violence-against-children/

United Nations Children's Fund. (2024). *Protecting children from violence and exploitation in relation to the digital environment: Policy brief*. https://www.unicef.org/documents/protecting-children-violence-and-exploitation-relation-digital-environment

Weeks, J. (2017). *Sex, politics and society: The regulation of sexuality since 1800* (4th ed.). Routledge. https://doi.org/10.4324/9781315161525

Wellington, M., Dew, A., Frawley, P., David, J., & O'Shea, A. (2025). Engaging people with intellectual disability in participatory action research: Co-developing sex educational resources. *Sexuality and Disability, 43*, Article 8. https://doi.org/10.1007/s11195-024-09886-3

**Contact emails:** cansu.bolatdinc@students.plymouth.ac.uk
bolatcansu96@gmail.com

# Bridging the Distance: Leveraging Communication Apps to Foster Collaboration and Reduce Isolation in Online Graduate Education

Gina Solano, Walden University, United States
Jennifer Miller, Lake Erie College, United States

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

Distance learning in graduate programs presents persistent challenges related to student isolation, diminished peer interaction, and the absence of spontaneous communication that characterizes face-to-face instruction. This qualitative study examined how the integration of mobile communication applications, specifically WhatsApp and Nectir, transformed collaborative dynamics among graduate students enrolled in an online Educational Technology Master's program at a university in Upstate New York. Drawing on open-ended questionnaire responses and observational data collected across three semesters from 65 participants, the study found that structured use of communication apps within small peer groups substantially reduced feelings of isolation, facilitated timely instructor-student exchanges, and cultivated a more cohesive learning community. Participants reported increased confidence in seeking help, stronger peer-to-peer relationships, and a more supportive course environment. These findings have practical implications for instructors seeking scalable strategies to strengthen collaboration and social presence in asynchronous online settings. The study is grounded in Anderson's (2011) Integrated Theory of Online Education, with particular emphasis on the community-centered and learner-centered dimensions of effective online pedagogy.

*Keywords:* distance learning, communication apps, collaboration, online graduate education, social presence, WhatsApp, Nectir

iafor
The International Academic Forum
www.iafor.org

## Introduction

Distance learning has become a permanent fixture in higher education, accelerated by the COVID-19 pandemic and driven by sustained demand for flexible, accessible graduate programming (Adedoyin & Soykan, 2023; Hamilton & Beagle, 2024). As colleges and universities expand their online offerings, a critical pedagogical challenge has emerged: how do instructors cultivate meaningful collaboration and reduce the social isolation that frequently undermines student success in asynchronous environments?

The problem is not simply one of logistics. Online learners regularly report feeling disconnected from both peers and instructors, a phenomenon that negatively affects motivation, academic performance, and course completion rates (Boling et al., 2012; Kaufmann & Vallade, 2022). Unlike a traditional classroom, where a student can turn to a neighbor, linger after class, or catch an instructor in the hallway, the distance learning environment typically constrains communication to formal channels (discussion boards, email, and periodic video sessions), none of which naturally support the informal, spontaneous exchanges that build community and trust.

This study investigates how the deliberate integration of mobile communication apps (WhatsApp and Nectir) within structured small-group configurations can address these gaps. Rather than treating communication apps as supplementary conveniences, the researchers examined them as pedagogical tools capable of reshaping the social architecture of online courses. The central argument is that when instructors embed app-based communication into course design thoughtfully, these tools do more than transmit information; they create the conditions for collaborative learning, peer support, and the kind of relational context that characterizes high-quality educational experience.

This is especially relevant in graduate education, where students often arrive with significant professional experience and benefit considerably from peer exchange, collaborative problem-solving, and a network of colleagues who share their field. For graduate students in programs such as Educational Technology, where applied, reflective, and collaborative competencies are central learning outcomes, communication infrastructure is not peripheral; it is foundational.

The following research questions guided this inquiry:

- What are the perceptions of graduate students who actively used WhatsApp or Nectir for communication in their online courses?
- In what ways did the use of these communication apps shape students' collaborative learning experiences?

## Context

The study was conducted within an online Educational Technology graduate program at a liberal arts university in Upstate New York. The program enrolls primarily K-12 educators who pursue their degrees while working full-time, necessitating a fully asynchronous course structure. Students are distributed across New York State, working in rural, suburban, and urban school districts, and complete all coursework through an institutional learning management system (LMS).

Two instructors in the program independently decided to supplement standard LMS communication tools with a dedicated communication app: one instructor used WhatsApp, and

the other used Nectir. Both instructors designed their use of the apps around the same structural approach: students were organized into small groups of three to five and given their own dedicated chat space within the app. Each group was assigned a name and a visual icon to build a sense of identity and to allow instructors to distinguish between peer conversations and messages directed at them.

A class-wide channel was maintained alongside the small groups, enabling instructors to broadcast announcements, share reminders, and respond to questions relevant to the entire cohort. Students were encouraged to use their small group space to ask questions, seek help on assignments, share resources, and simply check in with one another. This configuration was intended to lower the communicative threshold, giving students a low-stakes, informal space in which to interact, rather than relying exclusively on the more formal structures of discussion boards and email.

Both apps were available on mobile devices and desktops, which allowed instructors to maintain a visible presence in the course communication stream throughout the day without requiring scheduled office hours or synchronous check-ins.

## Communication Apps

Mobile communication applications have proliferated in both personal and professional contexts, and their educational potential has attracted increasing scholarly attention. For the purposes of this study, two applications were employed: WhatsApp and Nectir. Each offers distinct affordances worth examining separately.

### WhatsApp

WhatsApp is among the most widely adopted messaging platforms globally, with over 2 billion active users and an architecture designed for both individual and group communication (Dean, 2023). Its educational value has been documented across a range of contexts. Dahdal (2020) found that incorporating WhatsApp into assignment structures increased student motivation and active participation, while Chitanana (2023) demonstrated that the platform facilitated peer knowledge-sharing and community-building within university courses.

For online course communication, WhatsApp offers several pedagogically relevant features: real-time text messaging, voice notes, file and document sharing, read receipts that signal message acknowledgment, group chat functionality, and optional phone and video calls. The platform is free, requires no institutional subscription, and is accessible across devices. Its informal register, supported by emojis, GIFs, and multimedia, lowers barriers to participation, particularly for students who feel apprehensive about formal academic communication. WhatsApp's read receipts are particularly useful in a course context, as they provide both students and instructors with visible confirmation that messages have been seen, reducing anxiety about whether questions have been received.

### Nectir

Nectir is designed specifically for educational settings and addresses a key concern that some institutions have about consumer messaging platforms: student privacy. Rather than requiring a personal phone number, Nectir uses institutional email addresses to create accounts, making it compliant with student privacy regulations and more suitable for campus-wide adoption (Le,

2024). It can be integrated directly into an LMS and supports single sign-on capabilities, reducing friction in the login process (Nectir, n.d.).

While Nectir does not offer the full social feature set of WhatsApp, it compensates with tools tailored to academic use, including file sharing, course-specific channel organization, and a recently introduced AI assistant that can be customized by individual instructors. This AI tool draws on course syllabi to provide students with real-time responses to common questions, potentially reducing instructor workload while maintaining a responsive course environment (Ghai, 2024). Nectir also provides substantial implementation support through onboarding workshops, customer service, and instructional resources.

Both platforms support the kind of layered communication architecture used in this study: private peer-to-peer messaging, small group channels, and class-wide broadcasts. Both are accessible on smartphones and desktops, enabling instructors to remain present without being tethered to formal course spaces.

## Literature Review

### Framework: Integrated Theory of Online Education

This study is grounded in Anderson's (2011) Integrated Theory of Online Education, which synthesizes social-constructivist learning theory with the practical realities of independent online study. Anderson argues that effective online learning requires four interlocking conditions: learner-centeredness, knowledge-centeredness, assessment-centeredness, and community-centeredness. As Picciano (2017) emphasized, the essence of interaction among students, teachers, and content is central to constructivist theories of education.

Community-centeredness is of particular relevance here. Anderson (2011) maintained that students learn more effectively when they can work together, challenge one another's thinking, and collaboratively construct knowledge. This collaborative imperative is difficult to fulfill in asynchronous online environments, where the default communication tools (discussion boards and email) are not designed for spontaneous dialogue or the informal exchanges that deepen relationships and build trust. Communication apps, by virtue of their real-time responsiveness and conversational register, are better suited to support the community-centered dimension of online learning.

Learner-centeredness also informs this study, as the apps enable students to seek support precisely when they need it, whether that means clarifying an assignment the night before a deadline, checking in with a peer about a group project, or simply asking a quick question without drafting a formal email. This responsiveness is a structural affordance of the technology itself, not merely a product of instructor availability.

### Isolation and Its Consequences in Distance Learning

Isolation remains the most persistently documented challenge in online education. Kulikowski et al. (2021) identified inadequate communication as a primary source of difficulty for online students, citing both a lack of informational clarity and an absence of reciprocal exchange as contributing factors. Students who cannot easily ask questions or receive feedback in real time often disengage, not from lack of motivation, but from lack of access to the relational scaffolding that makes learning feel navigable.

The consequences of this isolation extend beyond student satisfaction. Rajabalee and Santally (2021) documented a clear relationship between peer and instructor interaction and academic achievement, finding that engagement was a significant predictor of learning outcomes. Pacansky-Brock et al. (2020) extended this finding to equity concerns, noting that students from underrepresented populations are disproportionately affected by communicative barriers in online courses. When asking for help requires navigating formal channels, composing polished emails, or waiting days for a response, those who already face systemic disadvantages are placed at further risk.

The instructor's role in mediating these dynamics is considerable. Tat et al. (2021) and Tarazi and Ruiz-Cecilia (2023) both found that the degree to which instructors intentionally create communicative opportunities (not merely making themselves available, but actively designing for interaction) shapes students' sense of belonging, engagement, and ultimately their success. Communication tools do not replace this intentionality; they enable it.

## Collaborative Learning in Online Environments

Online education has historically favored independent, self-paced learning structures over collaborative ones, partly because sustained collaboration is genuinely difficult to engineer asynchronously. Kalmar et al. (2022) described this as a paradox: the conditions of online learning (physical separation, asynchronous schedules, limited spontaneous contact) increase the need for social interaction even as they make it harder to achieve. Students engaged in complex, ill-structured tasks, such as design projects or case-based problem-solving, require vibrant learning communities to think through challenges, test ideas, and receive timely peer input.

Wang et al. (2022) confirmed that learner-to-learner interaction, alongside learner-instructor interaction, is a significant predictor of online engagement. Stone and Logan (2018) argued that informal communication channels, precisely the kind that platforms like WhatsApp enable, are particularly effective for creating the sense of connectedness that sustains engagement over the course of a semester. Cronjé and van Zyl (2022) similarly documented WhatsApp's effectiveness in building online learning communities, and Nasution and Munandar (2023) positioned it as a strategic tool for strengthening communication in higher education online courses.

## Instructor-Student Rapport and Communication

Research consistently identifies instructor-student rapport as a key determinant of online learning quality. Cole (2016) established a link between students' satisfaction with instructor communication and their overall course satisfaction. Kaufmann and Vallade (2022) found that students who perceived their instructors as communicatively engaged (responsive, approachable, and relationally invested) reported lower levels of loneliness and higher course satisfaction. Hansen-Brown et al. (2022) confirmed this dynamic in pandemic-era data, demonstrating that instructor friendliness and accessibility were associated with reduced student loneliness.

Meletiou-Mavrotheirs et al. (2021) found that 41% of online faculty used communication technologies daily and treated them as integral to course effectiveness. Yet Dolphy (2015) noted that many instructors in higher education remain hesitant to adopt informal communication tools. This hesitance may be addressed through institutional support, clear

guidelines about response expectations, and structured implementation approaches, all of which are addressed in this study's recommendations.

## Methods

This qualitative study employed a purposeful sampling design to examine the experiences of online graduate students who used communication apps in their courses. Following IRB approval (IRB #MODCR00000268), students enrolled in courses where instructors had implemented either WhatsApp or Nectir were invited to participate voluntarily. Data were gathered across three semesters, yielding a total sample of 65 participants.

### Participants

All participants were enrolled in a Master of Educational Technology Specialist program. They were K-12 teachers or permanent substitute teachers employed in public and charter schools across New York State, representing rural, suburban, and urban contexts. Most participants were in the first or second year of their graduate studies. This population offered a productive study context, as participants brought professional experience that informed their reflections on collaborative and communicative norms both in their classrooms and in their own graduate learning.

### Data Collection

The primary data instrument was an open-ended digital questionnaire administered via Google Forms and distributed through the LMS email tool. Open-ended questions were selected to allow for unanticipated insights and to capture the full range of participant perspectives without constraining responses to predetermined categories (Bachiochi & Weiner, 2008). Reminder emails were sent over three weeks to maximize the response rate.

In addition to the questionnaire, both instructors maintained observational journals throughout each semester. Journal entries documented patterns in student app use, the nature and frequency of interactions within the apps, and the instructors' interpretive impressions of how the communication tools were shaping course dynamics (Matteson, 2021). This observational layer was essential for capturing dimensions of the student experience that questionnaire responses might not fully reflect, including the texture and rhythm of day-to-day exchanges, instances of peer support, and moments when the app facilitated learning that would not have occurred through formal channels.

### Data Analysis

Data analysis was conducted solely by the researcher, who was also one of the two course instructors, using a manual, spreadsheet-based coding process (Google Sheets) grounded in open coding methodology (Creswell, 2012). All open-ended questionnaire responses were read in full and coded inductively, generating in vivo codes drawn directly from participants' own language. Codes were reviewed iteratively across multiple passes through the dataset and consolidated into broader categories aligned with the study's research questions.

Triangulation was achieved by cross-referencing questionnaire data with observational journal entries from both instructors, allowing the researcher to validate emerging themes across data

sources (Creswell & Guetterman, 2019). Theme saturation was reached when no new codes or patterns emerged from the data.

## Results

### Communication Accessibility and Reduced Anxiety

One of the most consistent findings across the dataset was the degree to which communication apps lowered the psychological and logistical barriers to asking for help. Students in online courses frequently reported feeling reluctant to send formal emails for what they perceived as minor questions, and many had experienced the frustration of waiting hours or days for responses when assignment deadlines were approaching. The apps directly addressed both of these problems.

Participant 24 articulated this clearly: “I feel that the tool is much more casual, which is nice because, at times, I just have a quick question and do not want to draft a formal email.” This casual register was not merely a matter of convenience; it represented a qualitative shift in students' willingness to seek help, with implications for their engagement and self-efficacy. Participant 19 described the cumulative benefit: “Considering I am in an entirely online program, the ability to message my professor and peers directly alleviate many opportunities for stress.”

The small group structure amplified these effects by creating a contained, lower-stakes space for communication. Participant 60 noted that the small group chat had made it easier to speak up: “It helps me to feel more comfortable having others to collaborate and speak with when completing assigned tasks.” Participant 59 drew an explicit comparison to the in-person classroom: “When you are in person, you can just turn and talk to the student next to you, but online, that is more difficult.” The small group format digitally replicated this dynamic, providing students with a nearby peer network they could consult without the formality of a class-wide announcement.

### Collaboration and Peer Interdependence

The communication apps were credited with significantly strengthening collaborative dimensions of course participation. Students reported using their group chats not only to ask logistical questions but to think through assignments together, share resources, provide peer feedback, and coordinate on group projects. This represented a meaningful shift from the transactional use of discussion boards to a more dialogic, reciprocal mode of engagement.

Participant 12 offered a notably thoughtful framing of this dynamic: “Online learning is not meant to be an entirely independent road. Rather, a support system, whether it be peers or the instructor, all should be pieces of the pie.” This perspective reflects a broader insight about online learning design: that the absence of collaborative infrastructure is itself a pedagogical choice, and that students notice and respond to whether instructors have invested in building connective tissue into the course.

Peer-to-peer relationships were valued not only instrumentally but intrinsically. Participant 22 commented: “I think it is important to be able to get in touch with my peers the most because we often have similar questions and are able to figure things out together.” This peer orientation is consistent with research on collaborative learning, which holds that co-construction of

knowledge produces deeper understanding than independent study alone (Anderson, 2011; Picciano, 2017).

Instructor observations confirmed that the apps extended participation beyond formal course activities. Students were engaging with one another and with course content in ways that did not appear in LMS analytics, having conversations about readings, offering encouragement ahead of due dates, and working through shared confusion collaboratively. This informal learning activity represents a substantive form of engagement that traditional course designs rarely support.

### Instructor-Student Relationships and Social Presence

A third major theme concerned the relationship between communication app use and instructor-student rapport. Students who interacted with their instructors through the app reported a qualitatively different sense of relationship than those limited to formal channels. Participant 18 reflected: “It is extremely important to create a bond and find a mentor. This connection helps students become better students and professors to become better professors.”

Participant 19 connected this relational dimension to professional identity:

> Some of the best relationships I formed during college were with my professors. They are the professionals in the field that I want to go into, so having access to them is an amazing resource. Using the tools that we have, I have felt engaged even though we are not a face-to-face program.

For graduate students in professional fields, instructor accessibility is not merely a comfort; it is an apprenticeship resource.

Participant 64 articulated how the relational context created by the app shaped the learning experience itself: getting to know the instructor better makes learning “less frustrating and more time efficient. It allows me to understand that they [the instructor] truly care about my experience in the class.” This perception of instructor care, expressed through informal, timely communication, appears to mediate the sense of isolation that otherwise characterizes asynchronous online learning.

### Limitations and Dissenting Perspectives

Not all participants responded positively to the communication apps. A minority expressed reluctance to download an additional application on their personal device. Some were uncomfortable with the instructor's presence in their small group chat. A few reported technical difficulties, and several indicated that the app had no discernible effect on their learning experience. One participant questioned the value of forming relationships in online courses altogether.

These dissenting perspectives, while proportionally small, carry practical significance. Instructors considering adopting communication apps should anticipate this range of responses and consider building flexibility into their implementation, for instance, offering alternative communication channels for students who prefer more bounded interaction, or making the app optional rather than required.

## Discussion

The findings of this study align with and extend existing research on communication, collaboration, and social presence in distance learning. Five interconnected themes merit particular discussion.

First, the structural role of the small group configuration deserves emphasis. The communication apps were not deployed as open, undifferentiated channels; they were organized around cohesive peer groups of three to five students. This design choice appears to have been pivotal. Research on group size in online collaborative learning suggests that smaller groups produce more equitable participation, stronger interpersonal bonds, and greater individual accountability (Anderson, 2011; Kalmar et al., 2022). By giving each student a small, named community to belong to, the instructors created the relational preconditions for effective collaboration before the first substantive interaction took place.

Second, the findings support the theoretical proposition that social presence, the degree to which participants perceive one another as real people in a mediated environment, is central to online learning quality. The communication apps enhanced social presence by enabling informal, spontaneous exchanges that are structurally absent from discussion boards and email. As Dailey-Hebert (2018) argued, maximizing interactivity in online courses requires moving beyond discussion boards to tools that support more naturalistic dialogue. The present study provides empirical grounding for this argument in a graduate context, with implications for instructional design practices.

Third, the equity dimensions of this intervention warrant attention. When students from underrepresented populations report that the app gave them the confidence to ask questions they would not otherwise have asked, this is not merely a convenience finding; it is an equity finding. Reducing communicative barriers in online courses has the potential to make these environments more equitable by lowering the threshold for help-seeking and broadening who participates actively in course discourse (Pacansky-Brock et al., 2020).

Fourth, these findings help clarify what these communication apps add beyond what the LMS already provides. Unlike institutional email or LMS-based discussion forums, WhatsApp and Nectir are accessible directly from students' personal phones and computers, independent of the university's login environment. This accessibility, rather than any feature unique to either platform, appears central to why students engaged with them: the apps met students where they already were, rather than requiring them to open a separate institutional system to communicate. This distinction positions the present findings less as evidence that a particular app is pedagogically superior and more as evidence that reducing the technical and psychological distance between students and their communication tools, regardless of platform, is what drives engagement. This contribution extends prior studies that examined WhatsApp or Nectir in isolation (e.g., Cronjé & van Zyl, 2022; Dahdal, 2020) by foregrounding accessibility itself, rather than any single platform's feature set, as the underlying mechanism.

Fifth, these findings should be interpreted with some caution regarding causal attribution. Participants in this study opted into or continued using instructor-implemented apps, and it is possible that students who were already more inclined to engage academically and socially were also the most likely to make active use of the communication apps and to respond favorably in the questionnaire. Instructor enthusiasm and consistent presence within the app, rather than the app itself, may also partly account for the reported increases in comfort and

connection; a less engaged instructor using the same tool might not produce comparable results. Disentangling the effects of the platform from the effects of instructor behavior and student self-selection remains an important task for the replication and longitudinal studies called for in the Limitations and Directions for Future Research section.

The broader implication of these findings is that communication apps, when thoughtfully integrated into online course design, function as more than supplementary tools. They constitute a form of instructional infrastructure: a structural investment in the relational conditions that enable learning. Instructors and program designers who treat communication architecture as a core pedagogical consideration, rather than an afterthought, are likely to produce meaningfully different outcomes for their students.

## Recommendations

Based on the findings of this study, the following recommendations are offered for instructors and program administrators considering the integration of communication apps in online graduate courses:

- Adopt a structured small-group model. Rather than creating a single class-wide channel, organize students into groups of three to five and give each group a dedicated communication space. This reduces social anxiety, promotes equitable participation, and builds peer relationships that sustain engagement over a full semester.
- Establish clear communication norms from the outset. Instructors should communicate expectations about response windows, appropriate use of the app, and the distinction between the app and formal course communication. Setting these parameters protects instructor boundaries while ensuring students know how and when to use the tool.
- Select the app to fit institutional context. WhatsApp is free, widely familiar, and feature-rich, making it an accessible starting point for individual instructors. Nectir or similar institution-facing platforms may be preferable where privacy regulations, LMS integration, or campus-wide consistency are priorities.
- Provide professional development and peer support for instructors. Many faculty are hesitant to adopt informal communication tools because they lack models for doing so effectively. Workshops, peer coaching, and curated examples of app-integrated course design can lower the barrier to adoption.
- Build evaluation into implementation. Instructors who adopt communication apps should collect systematic data on their use and effects, including both student experience data and observational records. This supports ongoing refinement and contributes to the growing evidence base on communication technology in online education.

## Limitations and Directions for Future Research

The primary limitation of this study is its geographic and institutional scope. Data were collected from a single university in Upstate New York, from two instructors within a single graduate program. While the sustained, multi-semester data collection strengthens the validity of the findings, generalizability to other institutional contexts, student populations, or disciplinary programs remains to be established.

Future research should pursue several lines of inquiry. Replication studies across multiple universities and program types would strengthen the evidence base and reveal whether the effects documented here are robust across diverse contexts. Longitudinal designs that follow

students across multiple courses, some with communication apps and some without, would enable more rigorous causal inference about the apps' effects on collaboration, engagement, and academic outcomes. Quantitative complements to this qualitative work, including validated instruments for measuring social presence and collaborative self-efficacy, would add rigor and comparability. Finally, research examining the differential effects of communication app integration for students from underrepresented populations would be particularly valuable, given the preliminary equity implications of the present findings.

## Conclusion

Distance learning, by its structural nature, separates students from the informal, relational dimensions of education that research consistently identifies as foundational to learning and belonging. Communication apps, strategically integrated into online course design, represent a practical, scalable, and low-cost means of addressing this gap. The findings of this study suggest that when graduate students have access to responsive, low-threshold channels for peer and instructor communication, they are more likely to ask for help, engage collaboratively, build meaningful relationships, and experience their online courses as genuine learning communities.

The pedagogical insight at the center of this study is both simple and consequential: the tools through which students communicate shape the quality of the relationships they form, and the quality of those relationships shapes the quality of their learning. Instructors who recognize communication infrastructure as a dimension of instructional design, worthy of the same deliberate attention as course content, assessments, and learning outcomes, are better positioned to create online environments that are not merely functional, but genuinely educative.

As online graduate programs continue to grow and diversify, the imperative to design for connection, not just for content, becomes increasingly important. Communication apps are one concrete, evidence-informed mechanism for doing so. This study offers both a rationale and a model for their use.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Claude, an AI assistant developed by Anthropic, was used in editing and formatting this manuscript. Its use was limited to correcting punctuation and grammar, verifying reference citation details against original published sources, and formatting the manuscript to conform to the target conference's submission guidelines. The author further declares that no AI or AI-assisted technology was used to generate, draft, or analyze the research content, findings, or interpretations presented in this manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Abu Hassan, H., Al-Rahmi, W. M., Yahya, N., Zakaria, M. A. Z. M., Kosnin, A. B. M., & Darwish, M. (2020). Development of a new model on utilizing online learning platforms to improve students' academic achievements and satisfaction. *International Journal of Educational Technology in Higher Education, 17*(38). https://doi.org/10.1186/s41239-020-00216-z

Adedoyin, O. B., & Soykan, E. (2023). Covid-19 pandemic and online learning: The challenges and opportunities. *Interactive learning environments, 31*(2), 863–875. https://psycnet.apa.org/doi/10.1080/10494820.2020.1813180

Ahern, T., Gooding, T., & Biedermann, N. (2024). Connect: A framework to enhance student connection to their course content, peers, and teaching staff in online learning environments. *Teaching & Learning in Nursing, 19*(2), 243–248. https://doi.org/10.1016/j.teln.2023.10.015

Alawamleh, M., Al-Twait, L. M., & Al-Saht, G. R. (2022). The effect of online learning on communication between instructors and students during Covid-19 pandemic. *Asian Education and Development Studies, 11*(2), 380–400. https://doi.org/10.1108/AEDS-06-2020-0131

Anderson, T. (2011). *The theory and practice of online learning* (2nd ed.). AU Press. https://doi.org/10.15215/aupress/9781897425084.01

Archambault, L., Leary, H., & Rice, K. (2022). Pillars of online pedagogy: A framework for teaching in online learning environments. *Educational Psychologist, 57*(3), 178–191. https://doi.org/10.1080/00461520.2022.2051513

Ashe, S. J., & Lopez, R. M. (2021). Communication in online learning – How important is it? *The Clearing House, 94*(1), 15–30. https://doi.org/10.1080/00098655.2020.1835795

Bachiochi, P. D., & Weiner, S. P. (2008). Qualitative data collection and analysis. In S. G. Rogelberg (Ed.), *Handbook of research methods in industrial and organizational psychology* (pp. 161–183). Wiley-Blackwell. https://doi.org/10.1002/9780470756669.ch8

Bellare, Y., Smith, A., Cochran, K., & Lopez, S. G. (2023). Motivations and barriers for adult learner achievement: Recommendations for institutions of higher education. *Adult Learning, 34*(1), 30–39. https://doi.org/10.1177/10451595211059574

Boling, E. C., Hough, M., Krinsky, H., Saleem, H., & Stevens, M. (2012). Cutting the distance in distance education: Perspectives on what promotes positive, online learning experiences. *The Internet and Higher Education, 15*(2), 118–126. https://doi.org/10.1016/j.iheduc.2011.11.006

Caskarulu, S., Maeda, Y., Richardson, J. C., & Lv, J. (2020). A meta-analysis addressing the relationship between teaching presence and students' satisfaction and learning. *Computers & Education, 157*(103966). https://doi.org/10.1016/j.compedu.2020.103966

Chitanana, L. (2023). What's up with the WhatsApp application in the classroom? *Education and Information Technologies, 29*, 10105–10125. https://doi.org/10.1007/s10639-023-12226-1

Chiu, T. K. F., Lin, T. J., & Lonka, K. (2021). Motivating online learning: The challenges of COVID-19 and beyond. *Asia-Pacific Education Research, 30*, 187–190. https://doi.org/10.1007/s40299-021-00566-w

Cole, A. W. (2016). Testing the impact of student preference for face-to-face communication on online course satisfaction. *Western Journal of Communication, 80*(5). http://dx.doi.org/10.1080/10570314.2016.1186824

Cole, A. W., Lennon, L., & Weber, N. L. (2021). Student perceptions of online active learning practices and online learning climate predict online course engagement. *Interactive Learning Environments, 29*(5), 866–880. https://doi.org/10.1080/10494820.2019.1619593

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson.

Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (6th ed.). Pearson.

Cronin, C. (2012). Memoing: An active tool in research. *The Reflective Researcher*. https://www.researchgate.net/publication/266500852

Cronjé, J. C., & van Zyl, I. (2022). WhatsApp as a tool for building a learning community. *The Electronic Journal of e-Learning, 20*(3), 296–312. https://doi.org/10.34190/ejel.20.3.2286

Dahdal, S. (2020). Using the WhatsApp social media application for active learning. *Journal of Educational Technology Systems, 49*(2). https://doi.org/10.1177/0047239520928307

Dailey-Hebert, A. (2018). Maximizing interactivity in online learning: Moving beyond discussion boards. *Journal of Educators Online, 15*(3). https://doi.org/10.9743/jeo.2018.15.3.8

Darby, F. (2024, February 9). *We still think online teaching isn't real teaching*. The Chronicle of Higher Education. https://www.chronicle.com/article/we-still-think-online-teaching-isnt-real-teaching

Dean, B. (2023, December 12). *WhatsApp user statistics 2024: How many people use WhatsApp?* Backlinko. https://backlinko.com/whatsapp-users

Dolphy, L. (2015). *Techniques and tools: Effectiveness of research-based pedagogical techniques through online learning tools with secondary students*. figshare. https://doi.org/10.6084/m9.figshare.1372472.v1

Ferri, F., Grifoni, P., & Guzzo, T. (2020). Online learning and emergency remote teaching: Opportunities and challenges in emergency situations. *Societies, 10*(4), 86–103. https://doi.org/10.3390/soc10040086

Ghai, K. (2024, May 1). *Case study: How the LAPU created the classroom of the future with Nectir AI.* Nectir. https://www.nectir.io/blog/case-study-how-lapu-created-the-classroom-of-the-future-with-nectir-ai

Hamilton, I., & Beagle, V. (2024, March 22). *By the numbers: The rise of online learning in the U.S.* Forbes Advisor. https://www.forbes.com/advisor/education/online-colleges/online-learning-stats/

Hansen-Brown, A. A., Sullivan, S., Jacobson, B., Holt, B., & Donovan, S. (2022). College students' belonging and loneliness in the context of remote online classes during the COVID-19 pandemic. *Online Learning, 26*(4), 323–346. https://doi.org/10.24059/olj.v26i4.3123

Hinze, A., Vanderschantz, N., Timpany, C., Cunningham, S. J., Saravani, S. J., & Wilkinson, C. (2023). A study of mobile app use for teaching and research in higher education. *Technology, Knowledge and Learning, 28*, 1271–1299. https://doi.org/10.1007/s10758-022-09599-6

Hollister, B., Nair, P., Hill-Lindsay, S., & Chukoskie, L. (2022). Engagement in online learning: Student attitudes and behavior during COVID-19. *Frontiers in Education, 7*(851019). https://doi.org/10.3389/feduc.2022.851019

Kalmar, E., Aarts, T., Bosman, E., Ford, C., de Kluijver, L., Beets, J., Veldkamp, L., Timmers, P., Besseling, D., Koopman, J., Fan, C., Berrevoets, E., Trotsenburg, M., Maton, L., van Remundt, J., Sari, E., Omar, L-W., Beinema, E., Winkel, R., & van der Sanden, M. (2022). The COVID-19 paradox of online collaborative education: When you cannot physically meet, you need more social interactions. *Heliyon, 8*(1). https://doi.org/10.1016/j.heliyon.2022.e08823

Kaufmann, R., & Vallade, J. I. (2022). Exploring connections in the online learning environment: Student perceptions of rapport, climate, and loneliness. *Interactive Learning Environments, 30*(10), 1794–1808. https://doi.org/10.1080/10494820.2020.1749670

Kordrostami, M., & Seitz, V. (2022). Faculty online competence and student affective engagement in online learning. *Marketing Education Review, 32*(3), 240–254. https://doi.org/10.1080/10528008.2021.1965891

Kulikowski, K., Przytula, S., & Sulkowski, L. (2021). Emergency forced pandemic e-learning – feedback from students for HEI management. *Open Learning, 36*(3), 245–262. https://doi.org/10.1080/02680513.2021.1942810

Le, V. (2024, June). *Nectir for schools - Deck*. Nectir. https://help.nectir.io/en/articles/6069935-nectir-for-schools-deck

Matteson, S. M. (2021). Chex Mix data analysis activity. *College Teaching, 69*(3), 121–125. https://doi.org/10.1080/87567555.2020.1843389

Meletiou-Mavrotheirs, M., Mavrou, K., & Rebelo, P. V. (2021). The role of learning and communication technologies in online courses' design and delivery: A cross-national study of faculty perceptions and practices. *Frontiers in Education, 6*, 1–17. https://doi.org/10.3389/feduc.2021.558676

Nasution, A. K. P., & Munandar, I. (2023). Trends, opportunities, and challenges of using WhatsApp in learning: A literature review. *Journal Simki Pedagogia, 6*(2), 531–544. https://doi.org/10.29407/jsp.v6i2.210

Nectir. (n.d.). *Welcome to the classroom of the future*. Nectir. https://www.nectir.io/

Pacansky-Brock, M., Smedshammer, M., & Vincent-Layton, K. (2020). Humanizing online teaching to equitize higher education. *Current Issues in Education, 21*(2), 1–21. https://cie.asu.edu/ojs/index.php/cieatasu/article/view/1905

Picciano, A. G. (2017). Theories and frameworks for online education: Seeking an integrated model. *Online Learning, 21*(3), 166–190. https://doi.org/10.24059/olj.v21i3.1225

Rajabalee, Y. B., & Santally, M. I. (2021). Learner satisfaction, engagement, and performance in an online module: Implications for institutional e-learning policy. *Education & Information Technologies, 26*(3), 2623–2656. https://doi.org/10.1007/s10639-020-10375-1

Salmon, G. (2013). E-*tivities: The key to active online learning*. Routledge. https://doi.org/10.4324/9780203074640

Scott, R. (2020, November 18). *Chat-based alternatives to online discussion forums*. Blog of the APA. https://blog.apaonline.org/2020/11/18/chat-based-alternatives-to-online-discussion-forums/

Singh, V., & Thurman, A. (2019). How many ways can we define online learning? A systematic literature review of definitions of online learning (1988–2018). *American Journal of Distance Education, 33*(4), 289–306. https://doi.org/10.1080/08923647.2019.1663082

Steele, J. P., Dyer, D., & Mandernach, B. J. (2023). You can't have it all: Faculty and student priorities in the online classroom. *International Journal for the Scholarship of Teaching & Learning, 17*(1), 1–9. https://doi.org/10.20429/ijsotl.2023.17108

Stone, S., & Logan, A. (2018). Exploring students' use of the social networking site WhatsApp to foster connectedness in the online learning experience. *Irish Journal of Technology Enhanced Learning Ireland, 3*(1), 42–55. https://doi.org/10.22554/ijtel.v3i1.28

Tarazi, A., & Ruiz-Cecilia, R. (2023). Students' perceptions towards the role of online teaching platforms in enhancing online engagement and academic performance levels in Palestinian higher education institutions. *Education Sciences, 13*(5), 449. https://doi.org/10.3390/educsci13050449

Tat, H. H., Nadarajah, Jie, T., Chan., Chin, T., Thoo., Ratan, A., Rafij, S., Hooi, K., & Beh. (2021). Do quality and satisfaction matter in online higher education? Turkish Online *Journal of Qualitative Inquiry, 12*(5), 4844–4852. https://www.tojqi.net/index.php/journal/article/view/5355

Vorlíček, R., Stašová, L., Hogenes, M., Janebová, E., & Průchová, Z. (2023). Issues and challenges of inclusion in distance teaching and learning. *Distance Education, 44*(2), 342–361. https://doi.org/10.1080/01587919.2023.2210528

Wang, Y., Cao, Y., Gong, S., Wang, Z., Li, N., & Ai, L. (2022). Interaction and learning engagement in online learning: The mediating roles of online learning self-efficacy and academic emotions. *Learning and Individual Differences, 94*. https://doi.org/10.1016/j.lindif.2022.102128

WhatsApp. (n.d.). *Feel like you're right there*. WhatsApp. https://www.whatsapp.com/stayconnected

Xu, B., Chen, N. S., & Chen, G. (2020). Effects of teacher role on student engagement in WeChat-Based online discussion learning. *Computers & Education, 157*, 103956. https://doi.org/10.1016/j.compedu.2020.103956

# The Homework as a Space for Dialogue in Curriculum Design

Barbara Bocchi, University of Trieste, Italy

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This paper explores the role of homework assignments within instructional design, conceptualizing them not merely as tools for assessment or consolidation, but as dialogic spaces where implicit conceptions of learning, teachers'design choices, and opportunities for students'action and interpretation intersect. Drawing on Bruner's view of culturally shared and often implicit educational assumptions, assignments are analysed as everyday practices through which such conceptions become observable. The theoretical framework combines Bernstein's distinction between visible and invisible pedagogy with Laurillard's Conversational framework, which conceptualizes learning as a dialogic cycle of description, action, observation, feedback, and adaptation. In this perspective, assignments are understood as potential feedback devices for both students and teachers, supporting reflective and adaptive instructional design. The study adopts a qualitative interpretative ethnographic approach and is currently being conducted in a comprehensive school in north-eastern Italy. Data includes audio recordings of classroom interactions, field notes, contextual descriptions, and institutional documents. The unit of analysis is the task-mediated design episode, defined as the situated set of practices, discourses, and artefacts through which assignments are designed, enacted, and reviewed. The study aims to examine how assignments are conceptualized and used, their role as feedback mechanisms, and the conditions that shape their dialogic potential.

*Keywords:* homework, mixed method, primary school

iafor
The International Academic Forum
www.iafor.org

## Introduction

Historically, homework has been one of the most pervasive and unquestioned pillars of global education systems. For decades, teachers, students and families have uncritically accepted the routine assignment of tasks to be carried out outside school hours, viewing it as the natural and inevitable extension of classroom teaching in terms of time and space (Cooper, 1989). Homework is defined in literature as the set of tasks prescribed by teachers for students to complete outside the formal school setting and without the teacher's direct and synchronous guidance (Yavich & Davidovitch, 2020). It has traditionally encompassed a wide range of formats: from highly repetitive worksheets and mnemonic grammar exercises to independent reading, solving algorithmic maths problems, and written compositions.

However, in recent years, practices relating to homework have undergone a profound epistemological and sociological transformation. They have moved away from their role as a passive backdrop to bureaucratic routines and have taken centre stage in heated pedagogical, scientific and media debates. This transition is closely linked to the evolution of parenting dynamics, the transformation of the school–family partnership and changing social expectations regarding childhood (Athienitis, 2026). As parenting models characterised by intense protection and constant cognitive and emotional support have gradually emerged, homework has shifted from being an individual and autonomous student duty to a shared and negotiated family endeavour. Consequently, the day-to-day management of homework has become a veritable "emotional barometer" within the home, capable of altering relational dynamics, generating performance anxiety, and triggering persistent intergenerational conflicts (Gill & Schlossman, 2004).

When analysed through the lens of micro-design in education, the decision to extend learning beyond the physical confines of the school building is not a neutral or purely organisational act, but rather a highly intentional pedagogical choice steeped in values. This task acts as a programmatic and operational interface between the formal space of the institution and the student's informal living environment. It transparently reflects the teacher's personal epistemology, their implicit models of learning, their conception of the subject, and their established professional routines.

Building on Jerome Bruner's well-known premise that everyday educational practices constitute an expression of the mind's and culture's "implicit theories," the setting of homework is, to all intents and purposes, a cultural text to be decoded (1996). From this perspective, homework becomes a genuine, observable mechanism capable of bringing to light what often remains unspoken within the lines of formal curriculum planning. Through its concrete form, in fact, the homework assignment reveals, first and foremost, whether the student is regarded as a mere passive recipient of procedures transmitted by the teacher or rather as an active agent and initiator of inquiry. At the same time, it reveals the underlying conception of the teaching process, showing whether learning is understood as a purely sequential accumulation of mnemonic automatisms or as a complex, negotiated and dialogical construction of meaning. Finally, the nature of the task allows us to understand the very concept of the home environment fostered by the school, demonstrating whether the home is viewed as a mere bureaucratic extension of the educational institution or as a genuine laboratory open to autonomous, reflective and prosocial experiences.

## Literature Review

The international scientific literature on homework reveals a structural polarisation between those who advocate the cognitive and self-regulatory functions of homework and its critics, who highlight its negative socio-emotional consequences and the risks of perpetuating social inequalities (Cooper et al., 2006).

### Arguments For and Against the Systematic Setting of Homework

Proponents of homework, mainly situated within the cognitivist educational psychology field, argue that it makes it possible to extend "time-on-task" (time devoted to learning) beyond the constraints of the school timetable. This extension is said to facilitate the consolidation of knowledge structures and procedural skills, as well as the development of transferable skills such as time and effort management, and autonomy and self-efficacy in study.

From this perspective, independent problem-solving outside the classroom is seen as a crucial testing ground for the transition towards self-regulated learning (SRL). When students tackle a task independently, they must engage in metacognitive processes such as monitoring, self-assessment, and managing environmental distractions.

Furthermore, constructivist pedagogy, drawing on Piaget's theoretical framework (1970), highlights how open-ended and challenging assignments can transform the home environment into a laboratory for applying the processes of assimilation and cognitive accommodation in contexts not controlled by the teacher.

### Behaviourist Degeneration and Instructional Drift

At the opposite extreme, critical literature denounces the persistence of neo-behaviourist approaches. These approaches reduce tasks to repetitive exercises, pre-formatted worksheets, decontextualised memorisation, and algorithmic automatisms. This type of task turns the home into a place where concepts taught in the classroom are passively reinforced and conditioned.

When a task lacks a clear pedagogical purpose, or when there is an imbalance between the clarity of the instructions and the cognitive complexity of the task, a phenomenon referred to in the literature as "instructional drift" occurs. Excessive rigidity, or conversely vague ambiguity, in the formulation of the task forces students to seek constant intervention from adults. What could potentially be a heuristic investigation degenerates into a sequence of externally guided tasks. This inhibits intrinsic motivation, generates cognitive avoidance, and reinforces states of learned helplessness (Hattie, 2009).

This educational drift is even more pronounced when tasks are used for disciplinary or punitive purposes, such as assigning extra work to discipline the class, or as negotiated rewards. Such practices are widely criticised by researchers as they erode spontaneous interest in knowledge.

### The Overload Effect and Meta-Analytic Evidence

Another criticism is the impact of homework on children's mental and physical health. An excessive amount of homework can have a negative impact on children's mental and physical health, depriving them of the time needed for free play, spontaneous socialisation, physical

activity and rest. This violates fundamental rights enshrined in international conventions on children's rights.

From a methodological and quantitative perspective, research reviews and meta-analyses conducted on large samples (including studies by Harris Cooper and John Hattie) unequivocally demonstrate that homework's cognitive effectiveness is strongly influenced by students'developmental and academic levels.

In primary school, the correlation between the amount of homework completed and academic achievement is essentially close to zero (negligible effect size). Younger children, whose executive functions and ability to inhibit distractions are still developing, derive minimal cognitive benefit from traditional homework. An excessive workload at this age is likely to lead only to an early aversion to schooling.

In lower and upper secondary school, as processes of self-regulation and formal abstraction develop, the impact becomes progressively more positive. At this stage, tasks encourage personal synthesis, critical thinking, and the integration of complex skills.

To avoid overwork, scientific literature recommends adopting quantitative guidelines, such as the 'rule of ten minutes' (ten minutes per day for each year of schooling), to balance time commitment with students' mental and physical wellbeing (Cooper et al., 2006).

**Amplification of Socio-Cultural Inequalities**

The most compelling critical argument is that homework acts as a powerful mechanism for reproducing and amplifying existing socio-cultural and economic inequalities. Since it is carried out in deeply heterogeneous and asymmetrical domestic environments, success in completing the work largely depends on the "cultural capital" (as defined by Pierre Bourdieu, 1986) that parents possess, as well as their availability to provide support or tutoring.

Students from disadvantaged socioeconomic backgrounds or from families of recent immigrants often lack suitable, quiet physical spaces, adequate internet connections, technological devices or adult figures capable of helping them interpret school texts. When schools set complex homework while assuming parental support, they turn a standardised activity into a mechanism that penalises the most vulnerable students and widens the achievement gap (Rønning, 2011).

**Educational Policies and International Perspectives**

In response to these challenges, international educational policy frameworks have adopted distinctly divergent strategies. The Finnish model, for instance, prioritises teachers' professional autonomy and minimises out-of-school tasks—particularly at primary level—by concentrating instead on making the most of classroom hours while safeguarding students' leisure time (Rønning, 2011). In a similar vein, recent legislative reforms in Poland have formally banned compulsory homework in the early years of primary school, rendering assignments in later years entirely optional and exempt from numerical grading with the explicit aim of alleviating childhood stress (Nolan, 2024). By contrast, Anglo-Saxon jurisdictions such as Ireland and Scotland maintain a decentralised, highly flexible approach by delegating the formulation of homework policies directly to individual schools; nevertheless, longitudinal research within these contexts reveals a widening divide between

students, whose attitudes towards homework grow increasingly negative with age, and parents, who frequently regard it as a primary benchmark of an institution's academic rigour (Children's School Lives, 2024). Finally, a range of innovative school networks have sought to shift the focus of curriculum design from rote factual acquisition towards social-emotional development by replacing traditional written tasks with prosocial activities, community inquiry projects, or "kindness diaries"—a movement echoed by comparative research across Southern and Eastern Europe, which underscores a rising disaffection with mechanical routines and calls for co-design frameworks that genuinely afford students a voice (Ivić, 2017).

### Domestic Relational Dynamics and the Role of Parents

Ethnographic studies of domestic interactions during homework reveal complex dynamics oscillating between constructive support and debilitating interference. Driven by a desire to ensure their children achieve high academic performance, parents often take over the problem-solving process from their children. This "overprotective" intervention deprives children of the opportunity to experience failure and develop self-regulation strategies (Montalbetti & Lisimberti, 2020).

Literature on instructional design suggests that parental involvement should shift from providing operational and content-related support to organising the learning environment, structuring routines, and offering emotional encouragement (light monitoring), while gradually reducing direct support (fading).

## Theoretical Framework

To move beyond the fruitless dichotomy between abolitionists and defenders of homework, this study is based on a conceptual framework that integrates three theoretical models from pedagogy and the sociology of education synergistically.

### Basil Bernstein: Visible and Invisible Pedagogies and Framing Rules

As an educational sociologist, Basil Bernstein (1975) developed a conceptual model of pedagogical discourse based on two key notions: classification, which denotes the strength of the boundaries between distinct academic disciplines; and framing, which measures the degree of control that teachers and learners exercise over selection, sequencing, pacing and criteria of learning.

Within this theoretical framework, Bernstein distinguishes between visible and invisible pedagogies. A visible pedagogy is characterised by explicit sequencing rules, transparent assessment criteria and robust teacher regulation. Traditional homework assignments, such as standardised worksheets, fill-in-the-blank exercises and single-answer drills, fall squarely within this category as their overarching objective is the accurate reproduction of a legitimised textual model. Conversely, an invisible pedagogy operates through an implicit instructional framework, shifting the primary focus onto students' internal processes of meaning-making and personal agency, with teacher control remaining subtle. Open-ended enquiries, self-directed research and the creation of personal artefacts align closely with this latter approach.

Applying the Bernsteinian framework to homework reveals a profound pedagogical paradox. When an educator assigns a task underpinned by an invisible pedagogy—such as instructing students to conduct open-ended research on a topic of their choice—without providing explicit

decoding rules, the assignment inadvertently creates a barrier for disadvantaged learners. Students who lack high levels of domestic cultural capital often find it difficult to work in these conditions as they lack the implicit sociolinguistic codes needed to convert a loosely structured activity into an academically valid performance. Consequently, to be equitable, a dialogic task must be skillfully designed.

Consequently, to operate equitably, a dialogic task must calibrate its framing masterfully: it must articulate its epistemic expectations and evaluative criteria explicitly, while preserving sufficient structural openness to allow for authentic, personal exploration.

**Diana Laurillard: The Conversational Framework**

Laurillard's Conversational Framework (2012) conceptualises formal learning as an iterative and dialogic cycle that engages both teachers and students continuously across two distinct yet interconnected cognitive domains: the conceptual (discursive and theoretical) and the practical (action-oriented and experiential). Laurillard's model articulates this dialogic cycle through five core stages. The first stage is description, in which the teacher introduces the core concept alongside its associated cognitive challenge. This is followed by the stage of student action, during which learners engage directly with the task by manipulating artefacts, solving problems or generating original texts. Next, during the observation stage, both teacher and student examine the outcomes of these actions and their broader epistemic ramifications. This leads directly to the feedback stage, which may be intrinsic, deriving from the context or artefact itself, or extrinsic, provided by the teacher or peers. The cycle then culminates in the adaptation stage, where both parties refine their conceptual models and subsequent actions in response to the insights gained.

Traditional homework assignments in prevalent educational practices severely disrupt this recursive cycle. The task is introduced in class during the description stage, completed at home in isolation during the action stage, and then marked in a directive, bureaucratic manner upon returning to school. This mechanical sequence eradicates reflective feedback and the adaptation phase. The present study applies the Conversational Framework to reconceptualise homework as a structural mechanism that restores and sustains the cycle's recursiveness across the home environment and the classroom.

**Robin Alexander: The Perspective of Dialogic Teaching**

Robin Alexander's Dialogic Teaching model (2018) positions structured classroom interaction as the fundamental driver of cognitive development and collective knowledge construction. Within this framework, a pedagogical approach is considered dialogic if it embodies the following five principles:

- Collectivity: teachers and students work together to address learning challenges within a community of inquiry.
- Reciprocity: this is characterised by mutual listening, sharing ideas and considering alternative viewpoints.
- Supportiveness: students can express ideas freely without fear of error or disparagement, and peers support one another to achieve a shared understanding.
- Cumulativity: participants build on each other's contributions to create coherent, cumulative arguments.
- Purposefulness: discourse is intentionally structured to achieve specific pedagogical and epistemic goals.

Historically, empirical research into dialogic teaching has focused almost exclusively on synchronous verbal interactions within the physical classroom. These are contrasted with the rigid, directive routines of the classic Initiation-Response-Evaluation (IRE) sequence. This paper proposes a fundamental theoretical shift by demonstrating that the dialogic quality of classroom review does not emerge spontaneously during interaction. Instead, it is structurally prefigured, conditioned and sustained by the design of the homework task itself. Far from representing an interruption to pedagogical dialogue, homework serves as its primary catalyst. When designed as open-ended, heuristic enquiry, it yields valuable cognitive artefacts, genuine questions, observations and constructive uncertainties that students can then bring back to school. These provide the essential raw material needed to initiate high-density dialogic co-construction.

## Research Methodology

### Research Objectives and Questions

The primary objective of this study is to formalise, operationalise, and empirically evaluate the Task-Mediated Design Episode (TMDE) as an analytical framework and tool for micro-design in teaching. The aim is to enhance the dialogic density of classroom interactions in primary schools. Specifically, the study will examine the influence of the operational structure of homework tasks (monological-closed vs. dialogical-open) on discussion dynamics, turn-taking structures, communicative repertoires, and epistemic network topologies during homework correction sessions.

The study was guided by three research questions (RQs):

RQ1 (Qualitative Mapping): How is the operational structure of the homework task (traditional monological vs dialogic open) reflected in the implementation of Alexander's five criteria (collectivity, reciprocity, supportiveness, cumulativity and purposefulness) during classroom review sessions?

RQ2 (Sequential and Temporal Analysis): What probabilistic variations occur in turn-taking patterns, particularly in the transition from the rigid IRE model to extended discourse chains, depending on the type of task completed at home beforehand?

RQ3 (Geometric-Spatial Modelling): How is the geometry of the class's epistemic networks restructured, and in what direction does the semantic centre of gravity of the interactional profiles shift when the discussion is mediated by tasks conceived as autonomous heuristic artefacts?

### Context, Sample and Empirical Corpus

This study takes a qualitative interpretative approach within an ethnographic framework, combining this with advanced techniques for the quantitative analysis of discursive interactions. The research was conducted over the course of one school year at a comprehensive school in north-eastern Italy.

The sample comprises:

- 12 primary school teachers (all tenured, with between five and 28 years' service);
- 240 primary school students spread across 12 Year 3, Year 4 and Year 5 classes.

Data were collected through high-fidelity audio recordings of homework correction and discussion sessions, non-participant ethnographic observations with field notes, semi-structured interviews with teachers and documentary analysis of assigned homework texts. The audio recordings were transcribed verbatim in full using the abbreviated Jefferson conversational notation system to ensure the anonymity of all participants.

### The Unit of Analysis: The Task-Mediated Design Episode (TMDE)

The definition of the TMDE lies at the heart of the methodological framework. The TMDE is not conceived as a simple worksheet or a single lesson, but rather as a trans-spatial, multi-code analytical unit that connects and articulates three distinct operational phases.

To objectively identify and extract the TMDE from the continuous flow of transcripts, a strict boundary rule was applied.

Start of the episode ($T_n$): this is formalised upon the appearance of an activation vector (a verbal cue or an explicit reference by the teacher or a student that anchors the discourse to the homework assignment).

Development of the episode ($T_{n+x}$): this comprises all consecutive turns in which the homework assignment, artefact or underlying concept is discussed, problematised and re-elaborated.

Closure of the episode ($T_m$): this is formalised when interaction regarding the task explicitly ceases and transitions to an unrelated teaching activity or definitive institutional summary.

### Coding Matrix and Quantitative Analysis Tools

The transcripts were coded using a two-level categorical matrix designed to classify the type of task (COMP_CHI: closed/traditional task vs. COMP_DIA: dialogic/open task) and the communicative moves and dialogic criteria evident in the turns of speech.

Two complementary techniques were used for advanced data processing:

Lag Sequential Analysis (LSA): This is used to calculate the conditional transition probabilities between different interactional codes at a specific time lag (Lag 1, Lag 2 and Lag 3). The statistical significance of the observed sequences compared with those expected by chance was verified by calculating the modified Allison–Liker z-score.

Epistemic Network Analysis (ENA): Used to model the structure and density of mathematical connections between codes within a moving co-occurrence window comprising three consecutive turns. ENA converts the co-occurrences of codes into vectors within a high-dimensional space, enabling a quantitative comparison of the interactional topology of traditional and dialogic tasks via t-tests on the centroids of the networks.

## Results of the Empirical Analysis

Applying Sequential Analysis and Epistemic Network Modelling in a coordinated way to a corpus of 135,000 words revealed discrepancies and structural transformations in the architecture of classroom interactions, depending on the assigned task.

**Excerpt A: Traditional/closed TMDE (COMP_CHI)–Subject: Mathematics**

In Excerpt A, which concerns marking a traditional exercise involving columnar calculations carried out at home, the interaction strictly adheres to the IRE triadic model.

> Teacher (T_1): Right, let's take a look at yesterday's worksheet. Exercise number two. Luca, read out the first calculation and tell me what it equals. [MOVE_TEST]
>
> Luca (T_2): Four hundred and twenty-five plus one hundred and thirty equals five hundred and fifty-five. [Student brief response]
>
> Teacher (T_3): That's right, well done! Five hundred and fifty-five. Has everyone written it like that? Good. Sara, do the second one. [MOVE_EVAL]
>
> Sara (T_4): Seven hundred and eighty minus two hundred and ten is five hundred and seventy. [Student brief response]
>
> Teacher (T_5): Correct. Correct your answers with a red pen if you've made a mistake. [MOVE_EVAL]

The teacher dominates the interaction entirely; turns of speech last an average of less than four seconds; feedback is immediate and definitive; and there is no horizontal exchange between peers. The homework is simply "validated" in a bureaucratic sense.

**Excerpt B: Dialogic/Open TMDE (COMP_DIA)–Subject: Science/Geography**

The teacher set the following heuristic homework task in Excerpt B: "Observe the shadows cast by an object of your choice at three different times of the day (morning, midday and afternoon), draw or photograph them, and write down your hypothesis as to why their length changes". The review in class takes on a completely different structure.

> Teacher (T_1): Let's look at the work that Elena has brought in. Elena has photographed the shadow of a chair in the garden. What do you notice when you compare the 9:00 am photo with the 12:00 pm one? [CRIT_PURP]
>
> Marco (T_2): At nine o'clock, the shadow is really long and looks like a giant. At midday, it's really short and is completely under the chair. [MOVE_ADD]
>
> Giulia (T_3): It was the same at my house too! I think it's because the sun is high in the sky, so the light shines straight down. [CRIT_CUM] [MOVE_EVI]
>
> Matteo (T_4): But the sun doesn't actually move – it's the Earth that's spinning! So, as the Earth spins, the angle of the light changes. [CRIT_REC] [MOVE_EVI]
>
> Teacher (T_5): Matteo has introduced an interesting idea about the angle of the light and the Earth's movement. Elena, what did you hypothesise in your observation diary? [CRIT_CUM] [CRIT_SUPP]

> Elena (T_6): I wrote that the ‘sun’ light bulb moves, but after listening to Matteo, I realised that we are the ones turning relative to the sun. [CRIT_CUM] [Conceptual Adaptation]

In this episode, the homework assignment becomes the starting point for a genuine scientific investigation carried out together. The children take turns to speak and offer complex explanations. They listen to one another, build on each other’s hypotheses and adjust their mental models in real time.

**Results of the Lag Sequential Analysis (LSA)**

The LSA provided a probabilistic formalisation of the observed differences in the transcripts, highlighting a break in the IRE pattern in dialogic tasks.

In the traditional sub-model (COMP_CHI), the probability of a teacher’s test question being followed by a student’s short answer at Lag 1 is 92%. Subsequently, the probability that a teacher’s directive evaluation follows the student’s response at Lag 2 is 88%. This probabilistic chain indicates a closed, rigid cycle that prevents divergence or spontaneous, in-depth exploration.

In contrast, within the DIALOGIC sub-model (COMP_DIA), displaying or citing a domestic artefact at Lag 1 is followed by learning discourse interventions comprising speculations, arguments, and hypotheses with a 64% probability. Furthermore, at Lags 2 and 3, the probability of observing horizontal expansion moves by peers, such as the addition of evidence or the formulation of counter-examples, is 42% and 38%, respectively. These values demonstrate quantitatively that the open-task structure breaks the teacher's monopoly on evaluation and extends the conversational chain.

**Results of the Epistemic Network Analysis (ENA)**

Modelling through ENA enabled the mapping of the geometric structure of interactions across the two contexts.

The COMP_CHI network exhibits a highly polarised, mono-axial and low-density structure. The dominant axis is formed by the edge connecting the test question to the teacher's evaluation, which has an edge weight of 0.85. Nodes related to reciprocity, cumulativity and students' spontaneous initiative occupy entirely peripheral and isolated positions.

In contrast, the COMP_DIA network exhibits a multi-polar, highly connected, balanced topology. Its central node is the home artefact, acting as a structural pivot that connects nodes representing reciprocity (CRIT_REC), cumulativity (CRIT_CUM), and justification or argumentation moves (MOVE_EVI).

A statistical comparison of the centroids of the two networks using an independent samples t-test confirmed that the differences in density and spatial configuration are highly significant, indicating a genuine restructuring of the classroom's cognitive geometry.

## Discussion

This study's empirical results contribute to the ongoing discourse on the role of homework in primary education by offering insights that move beyond the dichotomy of uncritical retention and radical abolition. By demonstrating how task structure influences turn-taking and the density of epistemic networks, this research shows that the cognitive and educational value of classroom review depends not only on the time dedicated to it, but also on the design of the Task-Mediated Design Episode.

### The Pedagogical Value of Heuristic and Open-Ended Homework

The data highlights the structural limitations inherent in assigning traditional, closed homework. When homework is reduced to mere executive drill or procedural completion, it creates a dual criticality: it isolates the student within the home environment without offering genuine cognitive challenges and reduces the review session upon returning to school to a bureaucratic-administrative check based on the IRE framework. Within this framework, homework loses its developmental value, becoming a source of frustration for families and passivity for students.

Conversely, dialogic and open-ended tasks demonstrate that, when framed as heuristic enquiry, authentic problem solving or the production of personal artefacts, homework can be a catalyst for dialogue. The instructional framework becomes more flexible, and learning becomes a process of shared negotiation. Under these conditions, Alexander's five criteria of dialogic teaching—collectivity, reciprocity, supportiveness, cumulativity and purposefulness—cease to be theoretical ideals and become observable practices in students' verbal interactions.

### Generative Artificial Intelligence as an Educational Co-designer

Recent technological transformations have opened up a key area of inquiry concerning the integration of generative artificial intelligence tools, such as large language models, into instructional design for homework tasks.

If used without foresight, technology could exacerbate a behaviourist approach. In this regressive paradigm, teachers may rely on AI to automatically generate repetitive exercise worksheets, while students may use it to passively generate answers, thereby eliminating cognitive effort and restricting assessment to syntactic checks alone.

However, when integrated into the TMDE model, generative AI can be a valuable co-designer for educators. Teachers can collaborate with AI to generate heuristic prompts that transform traditional exercises into open-ended enquiries. For example, a directive prompt such as "Solve ten addition problems with regrouping" could be replaced with an authentic task such as "Find three items at home whose combined price does not exceed £10, create an illustrated receipt, and explain your strategy for calculating the change to a peer." Furthermore, AI can facilitate scaffolding and personalisation tailored to different Zones of Proximal Development in the classroom, ensuring tasks are clear without being too easy.

Finally, AI can simulate realistic narrative contexts or ethical and scientific dilemmas for students to reflect upon with family members at home prior to collective classroom discussions. Through these transformative applications, technology supports teachers in constructing

meaningful bridges between informal domestic experiences and formal classroom discourse, rather than displacing student agency.

### Operational Implications and Guidelines for Professional Practice

Based on the empirical findings and operationalisation of the TMDE framework, specific operational guidelines have emerged for educators and school leaders seeking to reorient homework policies.

Crucially, this requires a shift in focus from quantity to quality, reframing homework as a means of generating artefacts. Rather than the routine execution of familiar procedures, assignments should encourage the creation of tangible outputs, such as drawings, written hypotheses, curated collections or authentic questions, to be shared and discussed in the classroom.

Consequently, the subsequent phase of classroom discussion must be formalised. Homework review should not be limited to mechanical worksheet checks or sequential readings of correct answers. Instead, it should occupy a dedicated space where teachers use dialogic teaching strategies such as rephrasing questions, treating errors as learning resources, requesting justifications and summarising students' contributions.

To ensure equity, teachers must be transparent about their pedagogical approach from a Bernsteinian perspective, explicitly articulating epistemic expectations to compensate for disparities in home cultural capital without presuming academic support from parents.

Finally, schools must actively transform the role of parents by forming alliances with families. By providing targeted guidance, institutions can foster indirect emotional support while discouraging intrusive intervention, thereby preventing homework from becoming a source of domestic conflict and punitive evaluation.

## Conclusions

This study demonstrates that homework is not a neutral element of educational routines, but rather an important micro-instructional design tool and trans-spatial dialogic space.

The introduction and empirical validation of the Task-Mediated Design Episode (TMDE) provide an integrated theoretical and methodological model that can describe, analyse and redesign the continuity of the learning experience between the home and the classroom. Through the coordinated use of Lag Sequential Analysis and Epistemic Network Analysis, the research shows that assigning open-ended, heuristic and dialogic tasks disrupts the rigid structure of the sequential IRE model. This transforms the classroom into an enquiry-based community in which students play an active, reflective and argumentative role.

In an era characterised by the rapid evolution of family structures and the advent of artificial intelligence technologies, reimagining homework necessitates moving beyond the unproductive dichotomy of either adhering to bureaucratic executive practices or uncritically abolishing it. Reconceptualising homework as an opportunity for dialogic design allows primary education to fulfil one of its most important functions: creating meaningful links between students' informal learning and formal institutional knowledge. This fosters a cognitive democracy based on listening, participation and shared critical thinking.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that ChatGPT was limited to and used for translation support from Italian to English, and language refinement. The research ideas, design, data, findings, interpretation, and academic responsibility remain with the author.

## References

Alexander, R. (2018). Developing dialogic teaching: genesis, process, trial. *Research Papers in Education,* pp. 1–38. https://doi.org/10.1080/02671522.2018.1481140

Bernstein, B. (1975). Class and Pedagogies: Visible and Invisible. *Studies in the Learning Sciences*, OECD, N.Y., pp. 1–39. ISBN 9264113347.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241–258). Greenwood Press.

Bruner, J. S. (1996). *The Culture of Education*. Harvard University Press.

Children's School Lives. (2024). *Longitudinal study on primary schooling in Ireland. UCD School of Education.* Children's School Lives (CSL) - UCD School of Education

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research*, *76*(1), 1–62.

Hattie, J. (2009). *Visible Learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Ivić, V. (2017). Učenička percepcija svrhe i korisnosti domaće zadaće. *Napredak* [Pupils' perception of the purpose and usefulness of homework. Progress]. 158(1-2), 85–102. https://doi.org/10.15516/cje.v26i1.5412

Laurillard, D. (2012). *Teaching as a Design Science: Building Pedagogical Patterns for Learning and Technology*. Routledge.

Nolan, L. (2024). *Why did Poland get rid of homework in primary schools?* RTE. https://www.rte.ie/news/2024/0421/1444630-po land-homework-primary-schools/

Piaget, J. (1970). *Psychologie et épistémologie. Pour une théorie de la connaissance* [Psychology and Epistemology. Towards a theory of knolodge]. Paris: Gonthier.

Rønning, M. (2011). Who benefits from homework assignments? *Economics of Education Review, 30*(1), 55–64. Who benefits from homework assignments?

Shaffer, D. W. (2017). *Quantitative Ethnography*. Cathcart Press.

**Contact email:** barbara.bocchi@units.it

# Coherence as a Coordination Problem: The Coherence Navigation Framework for Multi-agency Support in Neurodivergent Children's Education

Azibaolanari Okungbowa, Independent Scholar, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

When a child has additional support needs, their education does not happen in a classroom alone. It happens across a system: health appointments, local authority decisions, education plans, and third sector support, all supposed to connect. In Scotland, they frequently do not. Families fill the gaps. Children wait. This paper argues that fragmentation in multi-agency educational support is not primarily a resource problem, it is a coherence problem, and coherence can be built into systems, rather than left to families to construct alone under pressure. The Coherence Navigation Framework (CNF) drew on Phase 1 research with 50 formally recruited participants across five stakeholder groups in Scotland, generating over 85 interview and focus group contributions, plus a separate 50-response anonymous survey. Thematic analysis identified three conditions determining whether support pathways hold together for the children at their centre: shared navigation, where coordination is distributed across the system rather than absorbed quietly by families; visible navigation, where progress and gaps are transparent to all relevant parties; and supported navigation, where families receive structured guidance rather than repeated redirection. The CNF is an original contribution to understanding coordination failure in neurodivergent support systems, with implications for education, health, and social care policy. It responds directly to Scotland's Additional Support for Learning Review (McManus, 2026) and contributes to interdisciplinary literature on inclusive education, integrated support design, and the conditions under which educational outcomes improve for neurodivergent children. Coherence is not a soft ambition. It is a design requirement.

*Keywords:* additional support needs, Coherence Navigation Framework, information fragmentation, multi-agency coordination, neurodivergent children

iafor
The International Academic Forum
www.iafor.org

## Introduction

Multi-agency education systems are built on the assumption that information will move between services. In practice, it frequently does not. That gap shifts coordination work onto children, families, and schools in ways that directly shape educational decisions and outcomes.

When a neurodivergent child's support requires input from education, health, and social care, those services rarely share what they know. Each agency holds a partial picture. Each operates within its own referral systems, assessment frameworks, and timelines. The child moves between them. The information largely does not. What this produces is not merely an administrative inconvenience. It produces a structural condition in which the professionals closest to a child, teachers, learning support staff, school coordinators, are routinely required to make decisions without the full picture of that child's needs, history, or current circumstances.

What I was carrying into that room was the story of everything that had happened before. The referrals. The assessments. The decisions. The promises. The professional in front of me often had none of that. I was not just there to discuss my child. I was there to reconstruct the entire journey so far.

The moment a child's support depends on more than one service, coordination becomes an education issue, not just a service issue.

This has consequences that reach into classrooms. Schools plan provision, allocate resources, and make placement decisions on the basis of information that is incomplete not because professionals are careless but because the system was never designed to move information coherently. A child who presents one way to an education authority may be understood quite differently by the health service that holds their diagnosis, or by the social care team that knows their home circumstances. Each professional sees a portion. No single actor is routinely required to hold the whole.

Scotland's experience is instructive here. The Education (Additional Support for Learning) (Scotland) Act 2004, as amended, established statutory duties for multi-agency coordination around children with additional support needs. The Scottish Government's (2022) Getting It Right for Every Child framework placed further duties on services to work together around children's wellbeing. The Named Person provision, introduced in the Children and Young People (Scotland) Act 2014, went further still: it attempted to designate one professional as a coordination point for each child. The Supreme Court's ruling in Christian Institute v Lord Advocate (2016) ultimately limited the information-sharing provisions of that scheme. What the Named Person experience revealed, however, was not simply a legal problem. It revealed a design gap. Responsibility can be assigned by legislation. Coordination across information systems cannot be legislated into existence. It has to be built.

Scotland is not an outlier. The coordination problem it has attempted to address through legislation is a structural feature of multi-agency education and support systems in every jurisdiction where responsibility for children is distributed across agencies. Zurynski et al. (2025), in a scoping review of paediatric integrated care models across international contexts, found that even in formally integrated systems families described needing to coordinate the coordinators. Information barriers persisted. The picture no professional held in full was held instead by the family.

This paper draws on Phase 1 research with 50 participants across five stakeholder groups to propose the Coherence Navigation Framework (CNF) as a structural response to that problem. Its contribution to education research is precise. Schools cannot make consistently informed decisions about children when the information required for those decisions is fragmented across agencies. The CNF identifies the specific conditions under which that fragmentation is produced and the conditions under which it can be reduced.

Schools are being asked to make good decisions about children using information that is split across services that do not reliably share it. This paper names the structural conditions that determine whether that information ever arrives.

## Methodology

This paper draws on Phase 1 research conducted with 50 formally recruited participants across five stakeholder groups: 30 family members and 20 professionals drawn from NHS, local authority education, local authority social care, and third sector settings in Scotland. The qualitative dataset comprised over 85 interviews and focus group contributions, supported by 50 anonymous survey responses gathered separately through professional networks and community groups. The two datasets are distinct. Survey responses are used descriptively to contextualise qualitative findings. Percentage figures cited in this paper derive from the 30 formally recruited family participants, who completed a structured survey instrument as part of Phase 1 engagement.

The research used a mixed methods design. Qualitative data were analysed using reflexive thematic analysis (Braun & Clarke, 2006, 2022). Analysis was inductive. The framework presented in this paper emerged from patterns identified across all five stakeholder groups. It was not applied to the data from prior theory.

The lead researcher is also a parent of a child with additional support needs navigating Scotland's ASN system, a position that has directly informed both the lines of inquiry pursued and the interpretive lens applied throughout.

Phase 1 research was conducted by an independent researcher outside an institutional framework. No prospective ethics committee review was obtained. Informed consent was obtained from all participants. Data were anonymised before analysis and storage. Participation was voluntary. Data were handled in compliance with UK GDPR throughout. The researcher is a member of the Social Research Association and the research was conducted in accordance with SRA ethical guidelines.

## Findings

Analysis of the qualitative and survey data produced seven themes, organised into three clusters. Each cluster names a structural feature of multi-agency coordination failure as experienced by families and professionals in Scotland's ASN system. Together they provide the evidential foundation for the Coherence Navigation Framework.

### Cluster One: Information Does Not Move

Three themes emerged around the movement, or failure of movement, of information across agency boundaries.

The first theme identified information as structurally trapped within agencies rather than shared between them. Professionals across all four professional stakeholder groups described beginning assessments without access to records held by partner agencies. The child's history, diagnoses, and prior decisions were available in principle. In practice they did not travel. Each professional started again.

The second theme captured the burden this placed on family respondents. Ninety-seven percent of family respondents (n = 30, Phase 1 survey) reported difficulty navigating the system. Eighty-seven percent reported repeating the same information to professionals often or all the time. This was not occasional repetition. It was the condition under which every interaction took place.

The third theme identified the educational consequence directly. Schools described making provision decisions, placing children in support structures, and allocating learning support resources on the basis of information that was acknowledged by professionals to be incomplete. Decisions about what a child needed in the classroom were being made without knowledge of what health or social care had already established about that child.

More than one participant described the same experience: walking into a meeting and realising they were starting from the beginning again.

**Cluster Two: No One Holds the Map**

Two themes emerged around coordination responsibility.

The fourth theme found that coordination responsibility in multi-agency systems was implicitly assumed to be held by someone, but was not explicitly assigned to anyone. When asked who was responsible for ensuring information moved between agencies, professionals across stakeholder groups described uncertainty. Each assumed another actor held that responsibility. In practice no one consistently did.

What stood out was not that people refused responsibility. It was that everyone assumed someone else was already carrying it. When I asked who was coordinating, the answer was often uncertain. The role existed in theory more clearly than it existed in practice.

The fifth theme documented what filled that gap. Families assumed the coordination role by default. They attended meetings across agencies, carried documentation, ensured professionals in one setting knew what had been decided in another, and maintained the continuity of their child's support history across transitions. This role was sustained, invisible, and unacknowledged in formal system design. Schools frequently reported coordinating across agencies as well, absorbing a function for which they had neither the resource nor the mandate.

**Cluster Three: Support Ends Where It Is Most Needed**

Two themes addressed support across transitions.

The sixth theme found that professional support was consistently withdrawn at the precise points where the coordination demands on families were highest. When a child transitioned between stages of education, changed placement, or moved between services, the professional who had built understanding of that child closed the case. The incoming professional began

again. Families described these transition points as the moments when the system became most difficult to navigate.

Families described losing continuity. A case closed, a professional moved on, a service changed, and years of accumulated knowledge seemed to disappear with it. What followed was not a transition. It was a restart.

The seventh theme distinguished between signposting and staying. Families and professionals both identified a pattern in which available support took the form of directing families toward another service rather than remaining present to ensure handover produced continuity. Signposting discharged the professional's immediate responsibility. It did not produce coherent support for the family. The families who experienced something different described one thing consistently: a professional who stayed past the point where the system expected them to leave.

## The Coherence Navigation Framework

The findings across all three clusters point toward the same structural absence. Coordination is assumed. It is not designed. The Coherence Navigation Framework is proposed as a direct response to that gap.

The Coherence Navigation Framework identifies three structural conditions required for families to experience coherent support in multi-agency systems. It provides an analytical structure for assessing whether these conditions are present in contexts where responsibility, information, and decision-making are distributed across multiple actors, and where coordination is required but not explicitly designed.

The three conditions are Shared Navigation, Visible Navigation, and Supported Navigation.

Shared Navigation holds that coordination responsibility must be actively named and distributed. Someone must hold the map. Not assumed to hold it. Named, resourced, and accountable for holding it across agencies and across time.

Visible Navigation holds that pathway progress and decision history must be transparent to all parties, including families and the schools working with their children. The information a child's history contains must travel with the child, not remain with the agency that generated it.

Supported Navigation holds that families must receive structured, proactive guidance at points of need, and that professionals must remain present across transitions rather than withdrawing at the point where handover happens.

The three conditions are proposed as interdependent. Each amplifies the others. None operates at full effect alone. That claim is the subject of Phase 2 validation.

What the data showed was that none of these conditions failed on their own. When responsibility was unclear, visibility suffered. When visibility was absent, support became reactive. The same families and professionals were often describing all three problems at the same time. That was the clue that they were part of a single system rather than separate issues.

## Discussion

The three clusters of findings demonstrate precisely what the absence of each condition produces in educational settings.

Schools cannot make consistently informed decisions about children when the information required for those decisions is fragmented across agencies. The three clusters of findings demonstrate that this is not a failure of individual professionals. It is a failure of structural design.

Cluster One showed that information is not moving between agencies because no mechanism requires it to. Visible Navigation identifies the structural condition this absence represents and names what must be present for information to travel. When pathway progress and decision history are transparent across agencies, schools receive what they need to make informed decisions. When they are not, schools proceed without it.

Cluster Two showed that no one consistently holds coordination responsibility. Families absorb it by default. Schools absorb part of it too, without the resource or mandate to do so reliably. Shared Navigation addresses this directly. When coordination responsibility is named, the default does not fall to the family. It does not fall to the school either. It sits where it was always supposed to sit: within the system itself.

Cluster Three showed that support ends at transitions. The professional who understood the child left. The family navigated alone. Supported Navigation holds that continuity across transitions is a design requirement, not a matter of individual professional discretion.

The CNF's relevance to education is not peripheral. It is central. When Shared Navigation is absent, schools operate without a named coordination partner. When Visible Navigation is absent, teachers and learning support staff make decisions without the full picture. When Supported Navigation is absent, families arrive at key educational junctures, tribunal hearings (the formal process for challenging educational support and placement decisions), placement reviews, and transition planning meetings, carrying information no one else has held consistently on their behalf.

The Scottish experience of the Named Person provision is instructive here. In Scotland, the Named Person role was intended to ensure one identified professional held responsibility for a child's wellbeing across services. For an international audience, it represents the most sustained legislative attempt in any UK jurisdiction to address the coordination gap this research identifies. Its relevance to this paper is not political. It is analytical.

The CNF does not propose a new Named Person scheme or any equivalent legislative mandate. It proposes structural design standards that commissioners and practitioners can embed within existing frameworks without new legislation and without the information sharing provisions that led to the Supreme Court ruling in Christian Institute v Lord Advocate (2016). The CNF is not a repetition of Named Person. It is the structural response to what Named Person revealed about the limits of legislative mandating as a route to coordination.

What Named Person revealed was that assigning responsibility by law does not build the information architecture that coordination requires. The CNF proposes that the three conditions, Shared Navigation, Visible Navigation, and Supported Navigation, must be

designed into multi-agency systems regardless of the legislative framework in which those systems operate. They are not Scottish conditions. They describe what any multi-agency education and support system requires if it is to function coherently around a neurodivergent child.

The equity implication follows directly. Families with more resources, more confidence, and more capacity to advocate compensate more effectively when the system fails to coordinate. Families without those resources face the full structural weight of what the system never built. Schools serving children from families with fewer resources are therefore most exposed to the consequences of coordination failure. The CNF does not eliminate that inequity. It reduces the structural conditions that produce it.

One image stays with me. A parent sitting in yet another meeting with a folder full of reports, notes, emails, and assessments. The folder existed because the system had no equivalent. The parent had become the record, the navigator, and the coordinator all at once.

The Okungbowa and Lawani (in preparation) analysis of the Named Person gap in Scotland's ASN pathways provides further evidence of the resilience demands placed on families and informal leaders when structural coordination is absent. That analysis extends the argument developed here by examining how the gap created by incomplete legislative implementation was filled, and at what cost, by families, schools, and third sector workers operating without formal mandate or resource.

## Limitations

This study draws on 50 participants across a single national context, Scotland, and does not claim representativeness beyond that context. The findings are qualitative and context-specific, generated through reflexive thematic analysis rather than statistical generalisation. The Coherence Navigation Framework presented here has not yet been validated through the measurement instrument planned for Phase 2, and its three conditions should currently be read as an analytical structure grounded in this data rather than an empirically tested tool.

## Conclusion

Multi-agency education systems for neurodivergent children do not fail because the people in them stop caring. They fail because coherence was never designed in. This paper has argued that the failure is structural, that it is predictable, and that it can be named precisely enough to be addressed.

The three clusters of findings identify where structural coordination failure occurs. Information does not reliably move between agencies. No one consistently holds coordination responsibility. Support frequently withdraws at the transitions where it is most needed. These are not the products of individual shortcoming. They are the products of systems built on the assumption that coordination will happen without being explicitly designed.

The Coherence Navigation Framework proposes three structural conditions in response: Shared Navigation, Visible Navigation, and Supported Navigation. For education practitioners and commissioners, the CNF does not ask for more effort from already stretched professionals. It asks for different design. When coordination responsibility is named rather than assumed, when information follows the child rather than staying with the agency that generated it, and when

professional support is built to stay present across transitions, the school is no longer the last line of information management. It becomes what it was always meant to be: the place where education happens.

That claim rests not only on the research but on the researcher's own experience navigating the same system.

Yes, I believe it, because I have seen it happen, not everywhere and not consistently, but enough to know it is possible rather than aspirational. In my own experience of this work, including transitions I am still navigating for my own son, there have been moments when professionals had the information they needed, knew who was responsible, and stayed connected through a transition. In those moments the conversation changed. Less time was spent explaining the past. More time was spent thinking about the child and what happened next. I have also heard, just as clearly, from families for whom that was never true. Coherence is not the default. It is what happens when the conditions are deliberately in place.

For commissioners, the CNF provides a diagnostic structure. It identifies which of the three conditions is absent or partially present in any given multi-agency pathway. That is not a theoretical exercise. It is the difference between a school that receives a coherent picture of a child before a placement decision and a school that does not. It is the difference between a family that arrives at a tribunal hearing (the formal process for challenging educational support and placement decisions) carrying five years of documentation because no system held it, and a family that does not have to.

The CNF makes no claim that the structural conditions it identifies are sufficient on their own to eliminate inequity in educational outcomes for neurodivergent children. Resource, advocacy capacity, and family circumstance will continue to shape those outcomes. What the CNF addresses is the portion of that inequity that is structural in origin: the gap between what coordination was supposed to provide and what families and schools are left to provide instead.

This paper is one output from a research programme that is still developing. Phase 2 will test whether the three CNF conditions can be measured reliably, by developing an instrument designed to assess their presence or absence across different multi-agency systems and jurisdictions. The aim is a tool that practitioners and commissioners can use to identify where coordination is failing structurally, before that failure reaches the classroom, the placement decision, or the family that has run out of capacity to compensate for what the system will not hold.

Coordination failure in multi-agency systems supporting neurodivergent children is not inevitable. It is designed in. The CNF names what would have to be true for it to be designed out instead.

## Acknowledgements

The author wishes to thank Dr Adekunle Oke, accredited researcher with expertise in research design, methodology, and academic knowledge production across disciplines, for his contributions of substantive intellectual input and peer review to the research programme across multiple outputs. The author also thanks the 50 participants across five stakeholder

groups whose accounts form the foundation of this framework, and whose willingness to share their experiences made this research possible.

## Disclosure Statement

This paper is based on research presented at the European Conference on Education, UCL and SOAS University of London, July 2026, and forms part of a programme developing and testing the Coherence Navigation Framework. Related manuscripts are under peer review at the International Journal of Integrated Care and submitted to the British Journal of Social Work.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author used Claude, an AI language model developed by Anthropic, to assist with editing and improving the clarity of written expression in the preparation of this manuscript. All AI-assisted suggestions were reviewed, revised, and approved by the author. Claude was not used to generate research ideas, design methodology, collect or analyse data, develop the Coherence Navigation Framework, or produce any intellectual contribution. All research design, data collection, qualitative analysis, framework development, and conclusions are entirely the author's own independent work.

## References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1080/14780887.2006.9714573

Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. Sage.

Children and Young People (Scotland) Act 2014. asp 8. The Stationery Office.

Christian Institute and Others v Lord Advocate [2016] UKSC 51.

Education (Additional Support for Learning) (Scotland) Act 2004 (as amended by the Education (Additional Support for Learning) (Scotland) Act 2009). asp 4. The Stationery Office.

McManus, J. (2026). *Additional support for learning: review*. Scottish Government.

Okungbowa, A., & Lawani, A. (in preparation). *The Named Person Gap: Structural resilience demand and informal leadership in Scotland's ASN pathways*.

Scottish Government. (2022). *Getting it right for every child (GIRFEC): Policy statement*. https://www.gov.scot/publications/getting-right-child-girfec-policy-statement/

Scottish Government. (2023). *Pupils with complex additional support needs: research into provision*. https://www.gov.scot/publications/research-provision-pupils-complex-additional-support-needs-scotland/

Zurynski, Y., Hutchinson, K., Kang, Y., Vizheh, M., & de Groot, A. (2025). Family experiences of integrated care for children with medical complexity: A scoping review. *Child: Care, Health and Development, 51*(3), e70091. https://doi.org/10.1111/cch.70091

**Contact email:** mzzolaobi@yahoo.com

# Increasing Intraverbal Response Order Variability in a Child With Autism: Effects of Lag Schedule Combined With Differential Reinforcement

Masashi Kawamura, Hachioji City Utsukidai Elementary School, Japan

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Children with autism spectrum disorder (ASD) often develop rigid, repetitive language patterns. Although discrete-trial training effectively establishes intraverbal responses—verbal behavior controlled by another's verbal stimuli—it frequently produces stereotyped sequences. In the present case, a fixed emission order (e.g., carrot → cabbage → eggplant for "vegetables") led to blended utterances in which phonemes from adjacent words merged into a single unit, necessitating intervention. This study examined whether combining a Lag 1 schedule with differential reinforcement of different orders from the baseline sequence (DRDO) could increase response-order variability in a 12-year-old boy with ASD and severe intellectual disability. A multiple-baseline-across-behaviors design was employed across five semantic categories (fruits, vegetables, animals, insects, vehicles). Following baseline and instruction phases—during which instructions to vary response order produced no change—variability training was introduced sequentially. Two dependent variables were recorded: (A) the proportion of words emitted in a different order than the preceding trial, and (B) the proportion differing from the baseline sequence. Lag 1 alone produced minimal improvement. Adding DRDO—reinforcing sequences differing from the baseline order—substantially increased both measures in the first category (Non-overlap of All Pairs [NAP] = 1.00; Improvement Rate Difference [IRD] = 1.00 for Measure A; NAP = 0.94, IRD = 0.88 for Measure B). Gains generalized rapidly; later-trained categories reached criterion within one to three training trials, suggesting cross-category transfer. The classroom teacher responded positively on a 15-item social validity questionnaire. These findings suggest DRDO is an effective adjunct when instruction or Lag alone are insufficient and offer a classroom-feasible framework for sequencing interventions.

*Keywords:* autism spectrum disorder, intraverbal behavior, lag schedule of reinforcement, response variability, differential reinforcement

iafor
The International Academic Forum
www.iafor.org

## Introduction

Children with autism spectrum disorder (ASD) frequently demonstrate difficulties in acquiring functional verbal communication (American Psychiatric Association, 2013). According to Skinner (1957), verbal behavior is classified based on its functional relations with antecedent stimuli and consequences. Among the functional verbal operants, intraverbal behavior—defined as a verbal response evoked by a prior verbal stimulus without point-to-point correspondence—is considered particularly important because it plays a critical role in social interaction (Partington & Bailey, 1993). Intraverbal responding has been successfully established in children with ASD through discrete-trial training (DTT) using model prompts, prompt-delay procedures, and prompt fading (Finkel & Williams, 2002; Goldsmith et al., 2007; Ingvarsson & Hollobaugh, 2011; Partington & Bailey, 1993; Watkins et al., 1989). One form of intraverbal responding involves naming multiple members of a semantic category (e.g., saying "apple" and "banana" in response to "Name some fruits.").

A well-documented characteristic of ASD, however, is the presence of restricted and repetitive patterns of behavior (American Psychiatric Association, 2013). Consistent with this characteristic, children with ASD who have acquired verbal behavior through DTT with prompt-based procedures described above often emit fixed, stereotyped response patterns across trials (Sundberg & Partington, 1999). In the case of category-member naming, this stereotypy often manifests as emitting the same words in the same fixed order on every occasion. Such limited variability in verbal behavior may impair social interaction (Olin et al., 2020) and may also produce a variety of collateral problems, including reduced contact with reinforcement, failure to respond, and problem behavior (Contreras & Betz, 2016). These findings underscore the clinical importance of developing procedures that increase response variability.

One procedure that has been used to increase intraverbal response variability is instructive feedback (IF), in which the therapist provides verbal information indicating that alternative responses are acceptable (e.g., "You could also say it this way."). Anthony et al. (1996) and Carroll and Kodak (2015) reported that IF effectively increased variability in verbal behavior. However, IF relies on the learner's ability to follow verbal instructions, and it may therefore be ineffective for some individuals. As described below, IF was not effective for the participant in the present study.

A second approach involves the use of lag schedules of reinforcement (lag schedules), in which reinforcement is delivered only when the current response differs from those emitted during the preceding *n* trials. Under a Lag 3 schedule, for example, reinforcement is delivered only if the current response differs from each of the responses emitted during the three immediately preceding trials. The effectiveness of lag schedules in increasing intraverbal response variability has been demonstrated across multiple studies (e.g., Contreras & Betz, 2016; Koehler-Platten et al., 2013; Lee et al., 2002; Olin et al., 2020; Silbaugh & Falcomata, 2019; Susa & Schlinger, 2012; Wiskow et al., 2018). Among studies involving individuals with ASD, Lag 1 has been used most frequently because higher lag values (i.e., Lag 2 and above) do not consistently increase response variability (Baruni et al., 2014). Nevertheless, Lag 1 schedules may produce stereotyped alternating response patterns rather than genuinely variable responding (Yamagishi, 2000).

The present study evaluated a procedure for increasing variability in the order of intraverbal category-member responses in a child with ASD. A Lag 1 schedule was initially implemented;

however, this procedure failed to disrupt the participant's highly stereotyped response sequence, likely because of an extensive history of reinforcement for the baseline sequence and the tendency of Lag 1 schedules to generate alternating patterns of responding. Therefore, a combined procedure involving Lag 1 and differential reinforcement of different orders from the baseline sequence (DRDO) was implemented. Under DRDO, reinforcement was delivered only when category members were emitted in an order that differed from the participant's baseline sequence.

## Method

### Participant and Reinforcers

One 12-year-old boy with ASD and severe intellectual disability who attended a special education classroom participated in the study. At the time of school enrollment, his developmental quotient (DQ) score on the revised Kyoto Scale of Psychological Development was 32 (Postural-Motor: 49; Cognitive-Adaptive: 37; Language-Social: 17). Prior to the present study, the participant had acquired the target intraverbal behavior—naming members of semantic categories—through a discrete-trial procedure involving the presentation and systematic fading of model, textual, and pictorial prompts (Finkel & Williams, 2001; Goldsmith et al., 2007; Ingvarsson & Hollobaugh, 2011; Partington & Bailey, 1993; Watkins et al., 1989). Although he had successfully acquired this repertoire, maintenance required frequent review sessions. The present study was conducted as part of one such review period.

Verbal praise (e.g., “OK,” “Great”) was used as the reinforcer. Based on the participant's prior history of acquiring behavior under contingencies involving this form of praise, verbal praise had previously functioned as an effective reinforcer for the participant.

### Therapist and Setting

The participant's classroom teacher, who was also the first author, served as the therapist. All sessions were conducted in the participant's special education classroom, with the therapist and participant seated across from each other at a desk.

### Experimental Design and Preassessment

The target behavior consisted of intraverbal responses in which the participant named members of five semantic categories: fruits, vegetables, animals, insects, and vehicles. The discriminative stimulus was the therapist's verbal presentation of the category name (e.g., “Name some fruits”), and the response consisted of a sequence of intraverbal responses naming members of the target category (e.g., “melon, watermelon …”). A multiple baseline design across behaviors was used. The study consisted of three phases: baseline (BL)–instructive feedback (IF)–variability training.

All five categories entered the IF phase simultaneously. Variability training was introduced in a staggered fashion using a multiple baseline design across behaviors, with training initiated first for fruits, then vegetables, and finally for animals, insects, and vehicles concurrently.

A preassessment was conducted prior to the study to identify the participant's response repertoire. The therapist presented each category name and asked the participant to name as many members as possible. Each category was assessed a minimum of three times. Results

confirmed that the participant consistently emitted the same words in the same fixed sequence (hereafter referred to as the *original order*) across all assessment trials (see Table 1).

**Table 1**
*Response Repertoire and Original Response Order Established During Preassessment*

| Category | Original Response Order |
|---|---|
| Fruits | melon → watermelon → lemon → orange → cherry → peach → grape → apple → persimmon |
| Vegetables | green onion → potato → pumpkin → lotus root → corn → burdock → bell pepper → bamboo shoot → carrot → cucumber → cabbage → daikon → broccoli → cauliflower |
| Animals | rabbit → gorilla → deer → bear → cow → horse → panda → white bear → zebra → elephant → giraffe → lion → monkey → crane → chicken |
| Insects | dragonfly → butterfly → mosquito → praying mantis → bee → katydid → cicada → wasp → spider → stag beetle → ladybug → tent caterpillar |
| Vehicles | motorcycle → bus → helicopter → helicopter → fire engine → plane → ambulance → police car → bicycle → car → taxi |

*Note.* Arrow symbols (→) indicate the direction of sequential word emissions within a single trial.

## Dependent Variables and Interobserver Agreement

The therapist recorded the words emitted by the participant on a data sheet during each trial. Dependent variable data were collected for all BL and IF trials, and for test trials during variability training. Two response variability measures, adapted from Contreras and Betz (2016), were calculated for adjacent word pairs within each trial.

Measure A: *Percentage of transitions that differed from the immediately preceding trial.* For each adjacent word pair within a trial, the pair was scored as variable if the two words were emitted in an order different from the corresponding pair in the immediately preceding trial. When a test trial immediately followed a training trial, the immediately preceding training trial served as the comparison trial for this measure. For example, if the preceding trial included the transition lemon→watermelon, then the transition lemon→orange in the current trial was scored as variable.

Measure B: *Percentage of transitions that differed from the original order.* For each adjacent word pair within a trial, the pair was scored as variable if the two words were emitted in an order different from the corresponding pair in the preassessment original order. For example, because melon was followed by watermelon in the original order, the transition melon→lemon in a given trial was scored as variable.

The participant typically said “none” (*nai*, the Japanese word for “none”) when he could not produce additional words; this utterance was also recorded as a word and included in the calculation of both measures, because “none” functioned as part of the participant's response repertoire, indicating the termination of responding. In addition, repetitions of previously emitted words within a trial occurred occasionally; the order of subsequent productions of repeated words was included in the calculation of both measures.

A classroom teacher who was naïve to the study's hypotheses served as a secondary observer. This observer watched video recordings of the training sessions and recorded which of the 58 unique words in the participant's repertoire were emitted in a trial (each word was scored once per recording). Interobserver agreement (IOA) with the primary observer was calculated as point-by-point agreement and was 98.3%.

**Baseline**

The therapist presented the discriminative stimulus (e.g., “Name some fruits”) and the participant was asked to name category members. If the participant paused without emitting the next word, the therapist prompted additional responding by saying “What else?” until the participant said “none” or did not respond for approximately 2 s. Each word emission was followed by the delivery of the reinforcer. No model prompts were provided. This constituted one trial. The intertrial interval was approximately 5 s (a duration maintained across all subsequent phases). Two trials were conducted per category per session.

**Instructive Feedback**

Procedures were identical to baseline, except that following delivery of the reinforcer after each trial, the therapist provided verbal instructions specifying an alternative response order for the first two words of the participant's sequence on the next trial. For example, because the participant emitted the fruits category in the order melon→watermelon→lemon … during baseline, the therapist stated, “Next time, try saying lemon, grapes first.” Five trials were conducted per category per session.

**Variability Training**

Variability training included training trials and test trials. Two types of training trials were implemented:

1. Lag 1 + prompt: Reinforcement was delivered only when each transition differed from the corresponding transition in the immediately preceding trial. Model prompts were provided to occasion reinforcement-eligible responses.
2. DRDO + prompt: Reinforcement was delivered for each word emission only if the order of that word in relation to the preceding word differed from the original order established during preassessment. Model prompts were provided to occasion reinforcement-eligible responses.

Model prompts were initially delivered immediately and then faded using a progressive time delay procedure. Specifically, after two consecutive reinforced responses were emitted without errors, the prompt delay was increased to approximately 2 s, and once two consecutive reinforced responses were emitted at the 2-s delay, prompts were discontinued. However, additional prompts were occasionally provided outside of these criteria to maintain the participant's task engagement by increasing the rate of reinforcement; therefore, the prompt-fading procedure was implemented flexibly based on participant responding.

Transition from training trials to test trials occurred when the participant emitted three consecutive reinforcement-eligible responses without a prompt. In sessions in which both training types were implemented, DRDO + prompt training generally preceded Lag 1 + prompt training.

During the latter portion of variability training, the number of training trials was progressively reduced as data trends warranted. Some training trials were interrupted due to the need to attend to other students in the classroom who were working independently on academic tasks; the number of completed training trials is indicated in Figure 1.

Test trials followed the same procedure as baseline—no model prompts were provided—except that a Lag 1 reinforcement schedule was in effect: reinforcement was delivered for each word emission only if the order of that transition differed from the order emitted in the immediately preceding trial.

### Social Validity

A 15-item questionnaire modeled after the Intervention Rating Profile-15 (IRP-15; Witt & Elliott, 1985) was developed. A second teacher, who had provided direct instruction to the participant in contexts unrelated to the present study but was not involved in conducting the study, was given a brief description of the study and asked to rate each item on a 6-point Likert scale ranging from 1 (strongly disagree) to 6 (strongly agree).

### Ethical Consideration

Written informed consent was obtained from the participant's parents prior to study implementation and for the subsequent publication of results.

## Results

Figure 1 presents Measures A and B across all trials during the BL and IF phases and across test trials during the variability training phase. Numbers above data points indicate the number of training trials completed immediately prior to that test trial. Uncircled numbers indicate Lag 1 + prompt training trials, whereas circled numbers indicate DRDO + prompt training trials.

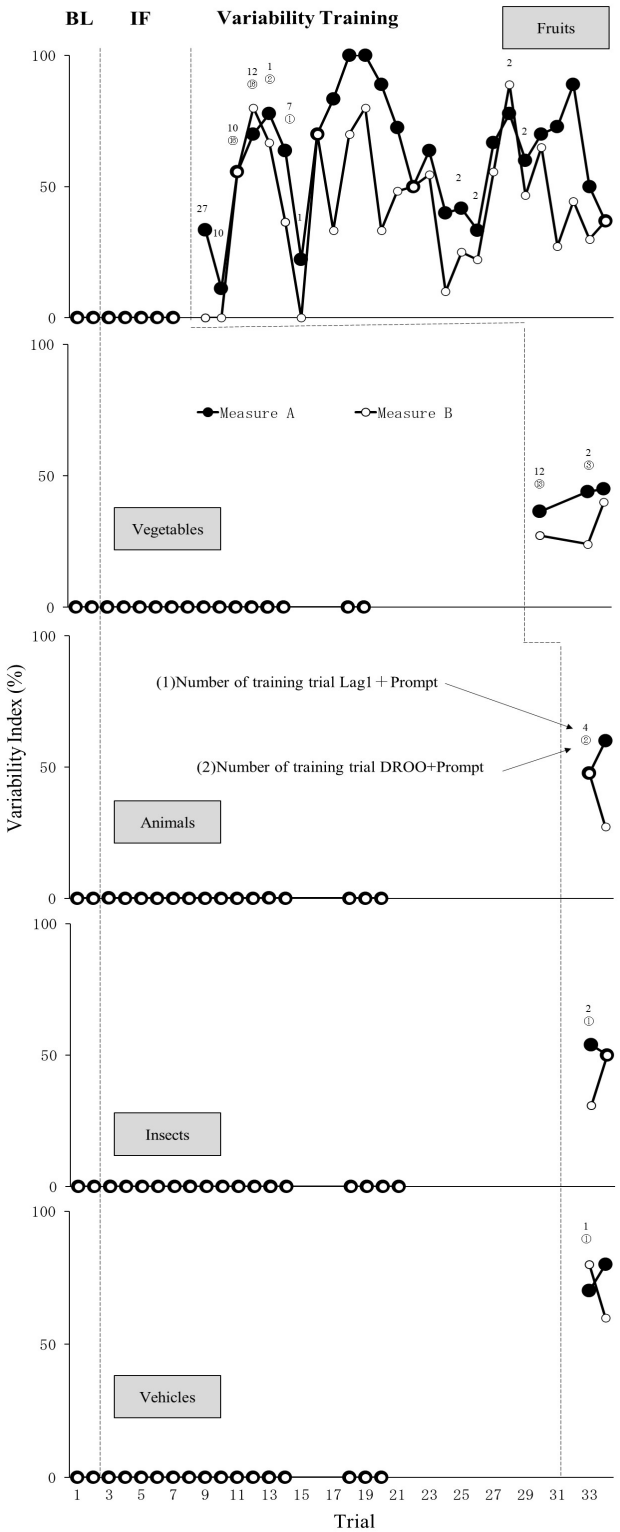
During the BL and IF phases, the participant consistently emitted category members in the original fixed order across all trials for all five categories, yielding values of 0% on both measures.

When variability training was introduced for the fruit category, Lag 1 + prompt training alone was initially implemented. The participant required 27 training trials to meet the criterion for transition to the first test trial, and subsequent test trial data showed no improvement or a decreasing trend on both measures. The participant began alternating between the original order and a single alternative order across trials. Based on this response pattern, DRDO + prompt training was introduced during the third set of training trials. Following the introduction of DRDO + prompt training, increases in both measures were observed. Similar improvements in both measures were observed across all remaining categories following the introduction of variability training.

The later categories (animals, insects, and vehicles) tended to meet the training criterion with fewer trials than the fruit category. In particular, the participant met the criterion for the vehicles category after only one training trial of each type (Lag 1 + prompt and DRDO + prompt), suggesting that fewer training trials were required for the later-trained categories.

**Figure 1**

*Changes in Response Order Variability Measures Across Baseline, Instructive Feedback, and Variability Training Phases*




*Note.* Filled circles (●) represent Measure A (percentage of transitions differing from the immediately preceding trial). Open circles (○) represent Measure B (percentage of transitions differing from the original preassessment order). Numbers above data points indicate completed training trials immediately prior to each test trial. Plain numbers indicate Lag 1 + prompt training trials; circled numbers indicate DRDO + prompt training trials. BL = baseline; IF = instructive feedback.

Effect sizes were calculated for the fruit category.[1] NAP and IRD (Parker et al., 2009) were calculated by comparing all BL and IF phase data points with all variability training phase data points. For Measure A, NAP = 1.00 and IRD = 1.00. For Measure B, NAP = 0.94 and IRD = 0.88.

The teacher who completed the social validity questionnaire provided predominantly positive ratings. Nearly all items were rated 5 (*agree*) or 6 (*strongly agree*). The sole exception was the item "This method of teaching category-member naming was consistent with the methods I typically use in my classroom," which received a rating of 4 (*slightly agree*).

## Discussion

The present study evaluated procedures for increasing variability in the order in which a child with ASD emitted category members during intraverbal responding. Although IF did not produce measurable changes in response order variability, variability training—combining a Lag 1 schedule with DRDO and model prompts—yielded increases in both dependent measures. Effect sizes for the fruit category exceeded established benchmarks: NAP values above 0.93 are classified as reflecting large or strong effects (Parker & Vannest, 2009), and IRD values above 0.70 to 0.75 are classified as reflecting large to very large effects (Parker et al., 2009).

The limited effectiveness of Lag 1 alone during the initial variability training of the fruit category is consistent with two known factors. First, the participant had an extensive training history with the original response order, which had been practiced repeatedly across prior review sessions, likely establishing a strong response preference for that order. Second, Lag 1 schedules have been reported to produce alternating response patterns rather than sustained variability (Yamagishi, 2000); such alternation was also observed in the present data. The addition of DRDO may have reduced the influence of the previously established response order and increased emission of alternative sequences.

Studies that have similarly targeted the order of multi-word emissions have reported that lag schedules (Contreras & Betz, 2016) and IF (Carroll & Kodak, 2015) can be effective. The present findings suggest that DRDO may be a useful supplementary procedure when these approaches alone are insufficient. It should be noted, however, that IF offers practical advantages: it is simple to implement (Carroll & Kodak, 2015), allows reinforcement to be delivered on every trial, and may thereby more easily sustain participant engagement. These findings suggest a possible decision sequence for practitioners. Practitioners might first attempt IF, then introduce a lag schedule if IF proves ineffective, and finally add DRDO if the lag schedule alone does not produce adequate variability. However, the present study did not include a component analysis comparing DRDO alone against the combined Lag 1 + DRDO procedure, nor was a design that would isolate the contribution of each component employed. Therefore, definitive conclusions regarding the independent or additive contributions of DRDO require further investigation.

[1] NAP is calculated by forming all possible pairs of data points between the baseline and intervention phases, assigning 1 point for each pair showing improvement in the expected direction, 0.5 for each tied pair, and 0 for each pair showing a change in the wrong direction, then dividing the sum by the total number of pairs. For full computational details, see Parker and Vannest (2009). IRD is calculated by removing data points in the baseline and intervention phases that overlap on the vertical axis and then calculating the proportion of remaining intervention-phase data points that changed in the expected direction. For full computational details, see Parker et al. (2009).

The tendency for the later-introduced categories to meet training criteria with fewer trials is noteworthy. The finding that the vehicles category—the last category to enter variability training—met criterion after only one training trial of each type raises the possibility that variability training may have produced generalized effects across semantic categories. However, it should also be noted that during the middle portion of fruit variability training, when training trials were eliminated, both measures showed a decreasing trend, suggesting that long-term maintenance without continued reinforcement may be a concern. The degree to which variability is maintained following training is an important question for future research.

The teacher respondent provided largely positive social validity ratings, consistent with prior research on the social acceptability of lag schedules (e.g., Contreras & Betz, 2016). The single item receiving a lower rating—consistency with typical classroom instruction—most likely reflects the fact that the procedures, including the lag schedule contingency, are not routinely employed in general or special education settings. This underscores the need for practitioner-friendly adaptations of behavioral procedures.

## Conclusions

Combining a Lag 1 schedule with DRDO and model prompts increased intraverbal response-order variability beyond the levels achieved with IF or a Lag 1 schedule alone, and these gains generalized rapidly across semantic categories. These findings suggest that DRDO is an effective adjunct when instructive feedback or a Lag 1 schedule alone are insufficient, and they offer a classroom-feasible framework that practitioners can adapt when sequencing variability interventions for learners with highly stereotyped intraverbal response patterns.

### Limitations

The following limitations of the present study should be acknowledged. First, only one participant was included, which limits the generalizability of findings. Second, although the participant's speech was highly intelligible, IOA data were collected on only a small subset of trials (approximately 3.5% of all trials), which constrains confidence in the measurement accuracy of the full dataset. Third, the criteria for delivering and fading prompts, as well as for reducing the number of training trials, were not operationally defined with precision, introducing potential for inconsistent implementation. Fourth, because a substantial number of sessions were devoted to the fruit category, the number of test trials completed for the remaining categories was too small to permit effect size calculations. Fifth, the procedures required one-to-one instruction, which requires that other students in the classroom can engage independently in academic tasks; thus, the approach may not be immediately applicable in all classroom contexts. Finally, the present study addressed only a basic form of intraverbal behavior—naming category members in response to a category cue. Whether these procedures would generalize to socially significant intraverbal behavior, such as conversational question–answering, remains unknown.

## References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Anthony, L., Wolery, M., Werts, M. G., Caldwell, N. K., & Snyder, E. D. (1996). Effects of daily probing on acquisition of instructive feedback responses. *Journal of Behavioral Education*, *6*(2), 111–133. https://doi.org/10.1007/BF02110228

Baruni, R. R., Rapp, J. T., Lipe, S. L., & Novotny, M. A. (2014). Using lag schedules to increase toy play variability for children with intellectual disabilities. *Behavioral Interventions, 29*(1), 21–35. https://doi.org/10.1002/bin.1377

Carroll, R. A., & Kodak, T. (2015). Using instructive feedback to increase response variability during intraverbal training for children with autism spectrum disorder. *The Analysis of Verbal Behavior, 31*(2), 183–199. https://doi.org/10.1007/s40616-015-0039-x

Contreras, B. P., & Betz, A. M. (2016). Using lag schedules to strengthen the intraverbal repertoires of children with autism. *Journal of Applied Behavior Analysis, 49*(1), 3–16. https://doi.org/10.1002/jaba.271

Finkel, A. S., & Williams, R. L. (2002). A comparison of textual and echoic prompts on the acquisition of intraverbal behavior in a six-year-old boy with autism. *The Analysis of Verbal Behavior, 18*, 61–70. https://doi.org/10.1007/BF03392971

Goldsmith, T. R., LeBlanc, L. A., & Sautter, R. A. (2007). Teaching intraverbal behavior to children with autism. *Research in Autism Spectrum Disorders, 1*(1), 1–13. https://doi.org/10.1016/j.rasd.2006.07.001

Ingvarsson, E. T., & Hollobaugh, T. (2011). A comparison of prompting tactics to establish intraverbals in children with autism. *Journal of Applied Behavior Analysis, 44*(3), 659–664. https://doi.org/10.1901/jaba.2011.44-659

Koehler-Platten, K., Grow, L. L., Schulze, K. A., & Bertone, T. (2013). Using a lag reinforcement schedule to increase phonemic variability in children with autism spectrum disorders. *The Analysis of Verbal Behavior, 29*(1), 71–83. https://doi.org/10.1007/BF03393125

Lee, R., McComas, J. J., & Jawor, J. (2002). The effects of differential and lag reinforcement schedules on varied verbal responding by individuals with autism. *Journal of Applied Behavior Analysis, 35*(4), 391–402. https://doi.org/10.1901/jaba.2002.35-391

Olin, J., Sonsky, A., & Howard, M. (2020). Using a lag schedule of reinforcement to increase response variability in children with autism spectrum disorders. *The Analysis of Verbal Behavior, 36*(2), 169–179. https://doi.org/10.1007/s40616-020-00129-y

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*(4), 357–367. https://doi.org/10.1016/j.beth.2008.10.006

Parker, R. I., Vannest, K. J., & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children, 75*(2), 135–150. https://doi.org/10.1177/001440290907500201

Partington, J. W., & Bailey, J. S. (1993). Teaching intraverbal behavior to preschool children. *The Analysis of Verbal Behavior, 11*, 9–18. https://doi.org/10.1007/BF03392883

Silbaugh, B. C., & Falcomata, T. S. (2019). Effects of a lag schedule with progressive time delay on sign mand variability in a boy with autism. *Behavior Analysis in Practice, 12*(1), 124–132. https://doi.org/10.1007/s40617-018-00273-x

Skinner, B. F. (1957). *Verbal behavior*. Appleton-Century-Crofts.

Sundberg, M. L., & Partington, J. W. (1999). *The need for both discrete trial and natural environment language training for children with autism.* In P. M. Ghezzi, W. L. Williams, & J. E. Carr (Eds.), *Autism: Behavior analytic perspectives* (pp. 139–156). Context Press.

Susa, C., & Schlinger, H. D. (2012). Using a lag schedule to increase variability of verbal responding in an individual with autism. *The Analysis of Verbal Behavior, 28*(1), 125–130. https://doi.org/10.1007/BF03393113

Watkins, C. L., Pack-Teixeira, L., & Howard, J. S. (1989). Teaching intraverbal behavior to severely retarded children. *The Analysis of Verbal Behavior, 7*, 69–81. https://doi.org/10.1007/BF03392838

Wiskow, K. M., Matter, A. L., & Donaldson, J. M. (2018). An evaluation of lag schedules and prompting methods to increase variability of naming category items in children with autism spectrum disorder. *The Analysis of Verbal Behavior, 34*(1–2), 100–123. https://doi.org/10.1007/s40616-018-0102-5

Witt, J. C., & Elliott, S. N. (1985). Acceptability of classroom intervention strategies. In T. R. Kratochwill (Ed.), *Advances in school psychology* (Vol. 4, pp. 251–288). Lawrence Erlbaum Associates.

Yamagishi, N. (2000). Effects of reinforcement contingencies on variability of serial responses in humans. *Japanese Journal of Behavior Analysis, 15*(2), 52–66. https://doi.org/10.24456/jjba.15.2_52

# Supporting Pasifika Learners in New Zealand Tertiary Education: Strengths-Based Pastoral Care Practices to Enhance Pasifika Learner Wellbeing and Educational Outcomes

Krishan Mani, Manukau Institute of Technology, New Zealand
Anamua Lole, Manukau Institute of Technology, New Zealand

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Pastoral support influences learner wellbeing but inconsistent support, limited transparency and confidentiality concerns can reduce its effectiveness. This study investigates effective pastoral care strategies for Pasifika learners, focusing on identifying core pastoral practices and understanding their relationship to learner wellbeing for enhancing educational outcomes. An online survey of Pasifika tertiary students received 20 responses from students enrolled in both part-time and full-time study in the areas of early childhood teaching and tertiary teaching. The research team used a pragmatic methodology combining Pasifika and Western perspectives, starting with a literature review and beginning talanoa (conversations) through an online survey. The findings reveal that effective pastoral care for Pasifika learners is fundamentally relational, holistic, and culturally grounded. Learners identified regular check-ins, academic guidance, and emotional support as the most valued forms of assistance. Feeling understood as Pasifika learners emerged as a central theme, with an emphasis on the importance of cultural recognition, respectful relationships, in environments free from discrimination. Pastoral support strongly influenced wellbeing, particularly emotional stability, motivation and mental health. Learners described how culturally safe pastoral practices foster a sense of belonging, strengthen inner mana (empowerment), and create confidence to engage in learning. The research demonstrates a clear connection between culturally responsive pastoral practices and improved wellbeing outcomes for Pasifika learners. It underscores the value of teachers' cultural capability, authentic relationship-building, and dedicated Pasifika support roles. The findings support the development of pastoral models that honour Pasifika values, identity, and community, contributing to enhanced learner wellbeing and educational outcomes.

*Keywords:* pastoral care, Pasifika learner, tertiary education, higher education, culturally responsive practice

iafor
The International Academic Forum
www.iafor.org

## Introduction

Pasifika learners in New Zealand's tertiary education sector face unique challenges that require culturally responsive pastoral care approaches. This study explores effective pastoral care strategies for Pasifika adult learners, synthesising research on culturally grounded support systems, relational pedagogies, and holistic wellbeing frameworks that enhance educational success for Pasifika learners. This research therefore examines strength-based approaches that honour Pacific cultural epistemologies while supporting learner wellbeing and achievement.

This project is a response to key findings from the Student Journey Playbook (Manukau Institute of Technology, 2021) which investigated the student journey at Manukau Institute of Technology. At the time, findings of the Student Journey Playbook identified Pasifika learners as one of the highest groups to withdraw from their enrolled program during their studies. The Student Journey Playbook found that Pasifika learners' experience impacted factors such as their attendance due to insufficient support from family and lecturers, inconsistent timetables, financial constraints and other commitments such as work demands. This research focuses on addressing these conflicting challenges and finding solutions to facilitate effective pastoral care for Pacific learners to complete their studies.

The project aims to identify:

- effective pastoral strategies from the Pacific learners' perspectives;
- Pacific centered approaches to meet the needs of Pacific learners while also seeking to improve their learning experience;
- strategies that will contribute to reducing student attrition rates and improve Pasifika education outcomes.

The project is divided into different stages:

- Stage 1: Define pastoral care
- Stage 2: Identifying best practices
- Stage 3: Identifying core components in pastoral care that practitioners should incorporate
- Stage 4: Exploring the correlation between pastoral care and health and wellbeing.
- Stage 5: Developing a set of guidelines highlighting effective pastoral strategies for Pacific learners

This report will be discussing key findings from stages 3 and 4 of the project by exploring the core components of effective pastoral care for Pasifika learners and examines the relationship between pastoral support, wellbeing, and learner success. During stage 3 and stage 4 of the project, data were gathered through surveys and qualitative feedback from Pasifika learners to identify commonly used pastoral practices and to explore how these practices influence health and wellbeing outcomes.

## Literature Review

The literature indicates educational systems need to be responsive to the multiple worlds of minority students, and increasingly policies and strategies in New Zealand have been adopted to promote the development of Pasifika learners. In the tertiary setting, improvements have occurred, although more is needed (McDonald & Lipine, 2012). The Education (Pastoral Care of Tertiary and International Learners) Code of Practice 2021 recognises the challenges

our student population can face and is in place to support the wellbeing and safety of learners studying with a New Zealand education provider (Sillifant, 2024). This regulatory framework has prompted institutions to examine how pastoral care can be delivered in culturally appropriate ways. In line with the Code of Practice, tertiary providers across New Zealand are looking at ways to support the mental health and wellbeing of adult learners, with recommendations focusing on taking a holistic approach to support both domestic and international students.

Pacific values are holistic both in understanding and action, and consider not only the physical but the emotional, mental, and spiritual aspects (Bentley-Gray, 2023). Pastoral care strategies must therefore address multiple dimensions of learner wellbeing rather than focusing narrowly on academic progress. Pastoral support was reported to strongly influence wellbeing, particularly emotional stability, motivation, and mental health. Learners described how culturally safe pastoral practices foster a sense of belonging, strengthen inner mana (empowerment), and create confidence to engage in learning.

## Methods

Talanoa, the Pacific practice of purposeful conversation, has emerged as a vital methodology for engaging Pasifika learners. The practice of Talanoa as a concept and a research tool in promoting Pacific knowledge systems and practices in tertiary education in Aotearoa New Zealand. The significance of recognising and acknowledging Pacific oral traditions will add value to and enhance Pacific ways of knowing and engagement in any context including that of Pacific learners in NZ tertiary settings.

Talanoa enhances meaningful engagement with learners as it provides a sound pedagogical tool in which to align conversations and draw responses that show what is significant in pastoral care for the Pacific learner. Talanoa methodology acknowledges the oral traditions of the Pacific voice which includes that of Pacific learners in the tertiary environment.

Talanoa is a widely used research methodology, often referenced to the initial and subsequent research work of Vaioleti (2006, 2016). Additionally, concepts of talanoa have been developed in Pacific research that highlight the types of engagement within the talanoa process such as malie (energy/uplifting) and mafana (warmth) (Manu'atu, 2000, as cited in Fa'avae et al., 2016). With the rise of technology, there is growing evidence to suggest that talanoa which is traditionally face to face, can be experienced in the virtual online space (Fairburn-Dunlop, 2014, as cited in Gremillion et al., 2021). This research has captured the use of talanoa in such a way that participants were able to respond through an online survey link.

### Demographic Information

- Age groups: 18 and above
- Area of study: tertiary teaching, early childhood teaching
- Level of study: diploma, degree–first year, second year, third year and above
- Pacific ethnic groups: Samoan, Tongan, Fijian, Cook Islands, Niuean

For this study, an online survey was used as the primary data collection method. The survey was distributed to adult Pasifika learners studying in teaching and adult education programmes at one tertiary institute in New Zealand. In total, 20 participants completed the

survey, providing valuable insights for the research. The survey consisted of 12 questions aligned with the project's aims and objectives.

## Results

### Effective Support Strategies

Participants rated the following 5 factors. All of the factors listed in Table 1 were rated as either important or very important. Notably, building positive relationships with staff and receiving academic support were among the highest-rated areas. The least rated factor—where only 45% of learners considered staff involving their family/aiga as very important, 10% indicated that it was not important. Although involving family is typically viewed as a key aspect of Pasifika values and norms, these findings suggest that from the learners' perspective, this is not seen as an essential component of pastoral support.

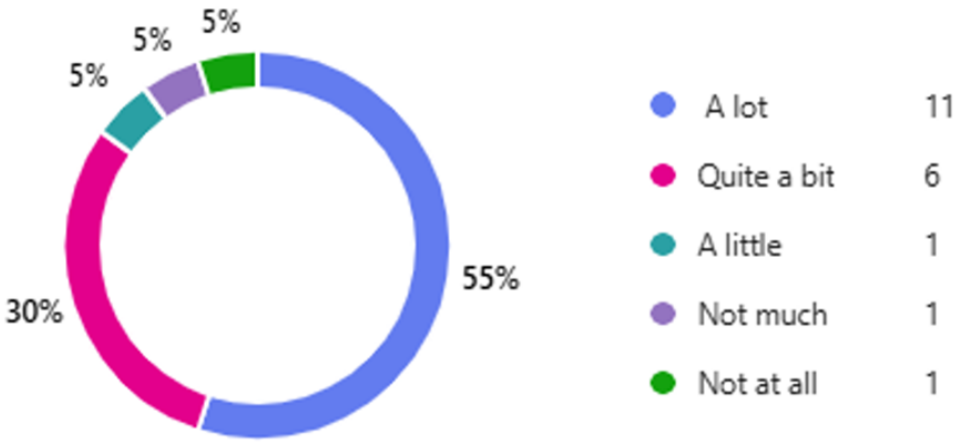
**Table 1**
*How Important Are the Following for You to Feel Supported?*

| | Very important | Important | Somewhat important | Not important |
|---|---|---|---|---|
| Building a positive relationship with staff | 90% | 5% | 5% | Nil |
| Staff understanding Pasifika culture | 70% | 25% | 5% | Nil |
| Regular check-ins about wellbeing | 55% | 30% | 15% | Nil |
| Staff involving your family/aiga when needed | 45% | 25% | 20% | 10% |
| Getting academic support | 80% | 20% | Nil | Nil |

### Pastoral Support Affects Wellbeing

The findings indicate that pastoral support plays a significant role in students' overall wellbeing. Many respondents (55%) reported that pastoral support affects their wellbeing "a lot," while a further 30% indicated it impacts them "quite a bit." This means that 85% of learners recognise pastoral support as having a strong positive influence on their wellbeing. However, 15% of respondents felt that pastoral support does not significantly affect their overall wellbeing. This suggests that while pastoral care is highly valued by most students, its impact may vary depending on individual circumstances, personal resilience, or the level of support experienced. Overall, the results highlight that pastoral support is a critical factor in supporting Pasifika learners' wellbeing, reinforcing the importance of consistent, meaningful, and culturally responsive support systems within the learning environment.

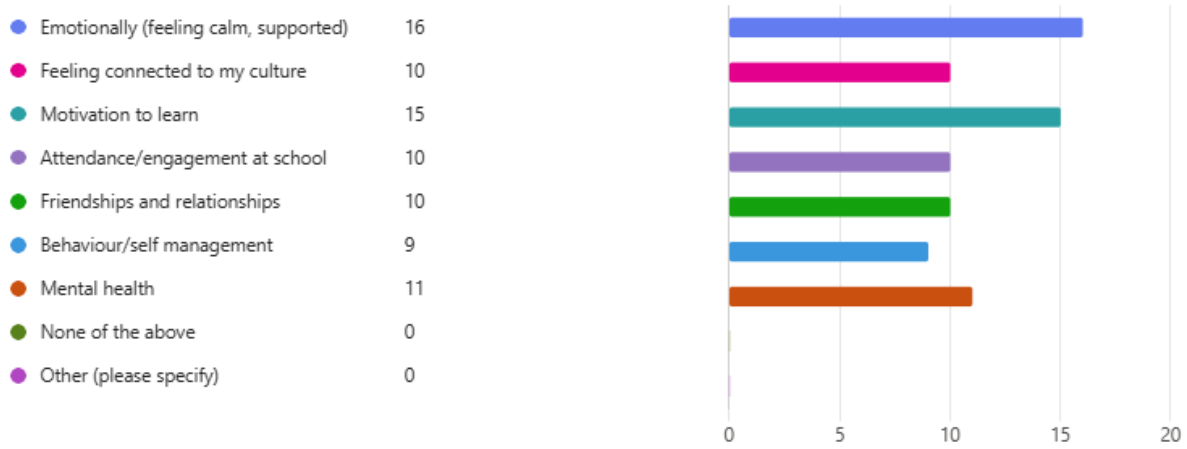
**Figure 1**
*How Much Do You Think Pastoral Support Affects Your Overall Wellbeing?*




## Pastoral Support and Academic Engagement

When asked which areas of wellbeing have improved because of pastoral support, the top 3 responses show that pastoral support has contributed most strongly to improvements in students' emotional wellbeing, motivation, and mental health. Students reported feeling an increased sense of calmness, support and reassurance, highlighting the importance of emotional care within pastoral systems. In addition, pastoral support has helped to strengthen students' motivation to learn, suggesting that when learners feel supported, they are more engaged and willing to succeed in their studies. Improvements in mental health were also identified, reinforcing the role of pastoral care in supporting students' overall wellbeing. The findings demonstrate that pastoral support plays a crucial role in fostering both wellbeing and academic engagement for Pasifika learners.

**Figure 2**
*Which Areas of Your Wellbeing Have Improved Because of Pastoral Support?*



## What Makes Pasifika Learners Feel Supported

Themes arising from the question "What makes you feel most supported as a Pasifika learner?" are:

### *A Sense of Belonging is Essential*

Feeling accepted, included, and free from discrimination helps students feel safe and confident in their learning environment.

### *Relationships With Lecturers are Key*

Support comes from lecturers who build genuine relationships, understand students, and check in on their wellbeing.

### *Teachers Need to be Approachable and Supportive Academically*

Students value being able to ask questions, receive clear guidance, and get help when needed.

### *Pasifika Values Strengthen Learning*

Family, community, cultural identity, and shared values play a big role in helping students succeed.

### *Institutional Support Also Helps*

Structured support like orientation and student services contributes, but is less impactful than relationships and culture.

## Discussion

Overall, the responses highlight that effective pastoral support for Pasifika learners is relational, culturally grounded, and holistic in nature. Students value support that recognises and affirms their cultural identity, as this strengthens their sense of belonging and ensures cultural safety. When cultural understanding is present, support feels meaningful; when it is absent, it can become a barrier.

A strong and consistent theme is the importance of relationships and connection. Learners feel most supported when teachers take time to build trust, listen genuinely, and engage with them beyond surface-level interactions. This relational approach aligns with Pasifika values of connection, community, and shared understanding.

Students also emphasise the need for holistic support, including emotional, social, spiritual, and mental wellbeing. Pastoral care is not seen as separate from academic success; rather, it is interconnected with students' overall wellbeing and ability to succeed. However, there are gaps in accessibility and consistency. While support is appreciated, it is not always regularly experienced, suggesting inequities in how pastoral care is delivered. In addition, trust and confidentiality are critical as without these, students may feel unsafe or reluctant to engage.

Another key insight is that effective pastoral care must include practical guidance and academic support, not just emotional encouragement. Students value clear communication, direction, and actionable advice alongside relational care. Additionally, pastoral support plays an important role in empowering students, helping them build confidence, strengthen their identity (mana), and contribute positively to their learning communities.

## Conclusion

The research demonstrates a clear connection between culturally responsive pastoral practices and improved wellbeing outcomes for Pasifika learners. Findings indicate that relational and culturally affirming pastoral strategies are central to enhancing Pasifika learner wellbeing and educational outcomes. Key forms of support identified included regular check-ins, academic assistance, and wellbeing support. Although majority of participants reported feeling well supported, learners consistently emphasised the importance of staff building genuine relationships, showing care through correct pronunciation of names, recognising specific island heritage, and demonstrating acceptance and respect for Pasifika identity.

Cultural safety, trust, and consistent communication were identified as critical enablers of effective pastoral care, while lack of confidentiality or limited follow-up reduced the perceived usefulness of pastoral support. Importantly, learners reported that pastoral support significantly affected their overall wellbeing, particularly in emotional wellbeing, motivation to learn, and mental health.

These findings support the development of pastoral models that honour Pasifika values, identity, and community, contributing to enhanced educational experiences and learner success. The findings suggest a strong correlation between pastoral care practices grounded in cultural responsiveness and improved wellbeing for Pasifika learners. This study highlights the need for institutional approaches that prioritise relational pedagogy, Pasifika-informed strengths-based pastoral care practices and regular proactive engagement from institute staff to enhance Pasifika Learner wellbeing and educational outcomes.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Grammarly and Copilot, AI-assisted writing software's, was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Bentley-Gray, D. (2023). *Talanoa: Pushing Boundaries to Promote Pacific Ways of Being in Aotearoa New Zealand Tertiary Education.* https://doi.org/10.34074/ocds.102

Fa‘avae, D., Jones, A., & Manu‘atu, L. (2016). Talanoa‘i ‘a e talanoa – talking about talanoa: Some dilemmas of a novice researcher. *AlterNative: An International Journal of Indigenous Peoples, 12*(2), 139–150. https://doi.org/10.20507/AlterNative.2016.12.2.3

Gremillion, H., Hallie, J., & Tominiko, F. (2021). The Scope of Talanoa Research Methodology: The place of research methods that are not rooted in Pasifika traditions. In Papoutsaki, E., & Shannon, M (Eds.), *Unitec Research Symposium Proceedings 2020* (pp. 39–53). ePress, Unitec New Zealand.

Manukau Institute of Technology. (2021). *MIT Student Journey Playbook MIT Tauira Haerenga Puka Takaro.*

McDonald, L., & Lipine, T. (2012). Pasifika education policy, research and voices: Students on the road to tertiary success. *New Zealand Annual Review of Education, 21,* 143–164.https://doi.org/10.26686/nzaroe.v21i0.4047

Sillifant, T. (2024). Mind the Gap! a University-Wide Approach to Finding Opportunities to Support Mental Health and Wellbeing in New Zealand Tertiary Institutions. *JANZSSA Journal of the Australian and New Zealand Student Services Association, 32*(2), 197–206. https://doi.org/10.30688/janzssa.2024-2-10

Vaioleti, T. M. (2006). Talanoa research methodology: A developing position on Pacific research. *Waikato Journal of Education, 12*, 21–34.

Vaioleti, T. M. (2016). Talatalanoa: Reframing talanoa to explore Pacific mental health. *Pacific Health Dialog, 21*(1), 15–22.

**Contact emails:** krishan.mani@manukau.ac.nz
anamua.lole@manukau.ac.nz

# Designing Career Readiness Through Industry-Embedded Capstone Pedagogy: A Curriculum Design Framework for Professional Identity Formation

Barbara Trippeer, University of North Texas, United States

The European Conference on Education, 2026
Official Conference Proceedings

**Abstract**

As the professional workplace landscape becomes increasingly complex and uncertain, educators are challenged to adapt to learning environments that move beyond traditional approaches focused primarily on isolated skills acquisition and employability credentials. This study examines an industry-embedded capstone model implemented across two upper-level fashion design courses involving forty students and proposes a curriculum design framework for professional identity formation. In this case study, the instructor deployed a scaffolded pedagogical approach for career readiness which activated students' individual agency through enculturation of discipline-specific professional learning communities. This model incorporated a sequence of integrated learning experiences, beginning with a discipline-specific career expo and progressing through several rounds of industry reviews, iterative project development, and a formal professional defense. These experiences allowed students to move beyond technical competence toward increasing participation in professional communities, fostering the development of professional identity. Drawing on student reflections and instructional observations, the study showcases how iterative engagement with industry professionals over several milestones influenced students' perceptions of their own professional development. The findings indicate that scaffolded industry participation supported key dimensions of professional development, including communication, critical thinking, professional judgement, reflective practice, confidence, and agency. Interpreted through experiential learning, communities of practice, and professional identity formation perspectives, the findings suggest that career readiness emerges not from isolated skills training but from intentionally designed learning architectures that support authentic participation, reflection, and professional socialization. The paper concludes by proposing a transferable curriculum design framework for fostering career readiness across disciplinary contexts.

*Keywords:* career readiness, professional identity formation, curriculum design, communities of practice, learning architectures, experiential learning

iafor
The International Academic Forum
www.iafor.org

## Introduction

Regardless of discipline or locality, there are several challenges for educators seeking to prepare students for successful entry into the workforce, not the least of which include economic uncertainties and destabilization of workflows due to the onset of artificial intelligence. Despite widespread attention to employability and workforce preparation, graduates continue to struggle with key dimensions of professionalism, such as professional judgement, communication, adaptability, critical thinking and decision-making in workplace environments (Becker, 2017; Park et al., 2026). If students can demonstrate technical competence, why are so many still unprepared for professional practice? Career readiness may involve more than acquiring professional competencies. It may require transforming the assumptions, values, and perspectives through which students understand professional practice.

Many instructions of higher learning actively seek to address these challenges through enhanced learning experiences, such as internships, practicums, service learning, or industry-embedded projects. Yet, while all of these have shown to be somewhat beneficial, participation alone does not automatically produce transformative learning. Although such experiences can provide valuable exposure to professional practice, learning emerges not simply from experience itself but from how experiences are intentionally designed, scaffolded, and integrated into curriculum (Billett, 2015; Tummons, 2014). Through participation in authentic professional communities, engagement with critique, reflective practice, and increasing responsibility for independent decision-making, students begin to internalize professional norms, values, and ways of thinking (Cranton, 1997). From this perspective, career readiness may be understood not merely as the acquisition of competencies but as a process of becoming a professional within a larger community of practice (Billett, 2015; Lave & Wenger, 1991; Wenger, 1998).

To support this process of professional becoming, higher education may require a shift away from curriculum models which are focused primarily on discrete skills acquisition instead towards intentionally designed learning architectures that foster professional identity formation (Foster, 1997; Mezirow, 1997). Tummons' (2014) work conceptualizes learning architectures as intentionally assembled systems that include all the people, places, and things (literature, materials, resources, classroom spaces, activities, etc.) that collectively shape learning (Tummons, 2014). By intentionally creating environments that provide legitimate opportunities for professional participation, educators establish conditions in which students not only become familiar with professional expectations but also begin to see themselves as members of professional communities (Lave & Wenger, 1991; Wenger-Trayner et al., 2014). Professional identity formation therefore serves as a critical bridge through which educational experiences become career readiness outcomes. From this perspective, career readiness emerges not from isolated activities, but from intentionally designed learning environments that support participation, identity development, and professional socialization (Billett, 2015).

This study explores how intentionally designed learning architectures support professional identity formation through participation, reflection, critique, dialogue, and authentic professional engagement. (Shor & Freire, 1987). Additionally, most learning environments are designed to reduce or control uncertainty, to maximize the output or to allow instructors to be able to assess specific learning outcomes or skills (Faerm, 2023). However, real-world professional environments are rarely static. Professional practice routinely involves ambiguity, competing variables, incomplete information, and a multiplicity of solutions. Hodges' (2019)

findings suggest that uncertainty may not be an obstacle to learning, but rather a necessary attribute required to achieve transformation, a state whereby professional judgement emerges during repeated encounters with uncertainty (Hodges, 2019). Uncertainty creates conditions for critical reflection and perspective transformation.

What distinguishes this model is not the industry engagement itself, but the intentional sequencing of increasingly critical professional experiences across multiple stages of development. This study examines an industry-embedded capstone model as an example of intentional curriculum design and explores how scaffolded participation in professional communities, structured critique, reflective practice, and repeated engagement with uncertainty contribute to professional identity formation and career readiness. The case study example strategically employed iterative cycles of engagement with industry professionals as a means through which to activate students' emerging professional identities. While many institutions have explored the use of similar programming, such as guest speakers, networking events, internships, or industry projects, which makes this example different is its application of a specific developmental sequencing while increasing the stakes of the level of experience.

This example demonstrates the application of not only Kolb's (2015) active learning cycle of reflection, but also the inclusion of critique as a mechanism for professional formation (Kolb, 2015; Schön, 1983). Mezirow argues that discourse and critical reflection allow learners to examine assumptions, consider alternative perspectives, and transform existing frames of reference (Mezirow, 1997). In this context, critique functions not only as professional socialization but also as a mechanism for transformative learning. Through dialogue, critique, and critical reflection, students are challenged to examine assumptions, justify decisions, negotiate competing perspectives, communicate reasoning, and internalize disciplinary norms while refining their understanding of professional practice (Grabove, 1997; Mezirow, 1997).

This paper advances the position that career readiness is an emergent outcome of intentionally designed learning architectures that support participation in professional communities, professional identity formation, and the development of judgement, agency, and reflective practice.

## Literature Review

### Beyond Employability Skills

Career readiness has become a pivotal focus in higher education, as institutions aim to equip graduates for an increasingly complex and uncertain professional landscape (Haynes, 2026). Contemporary literature often emphasizes employability skills and technical competences, however, feedback from industry professionals indicates significant gaps in graduates' adaptability, communication skills, critical thinking, and decision- making (Adu et al., 2025; Gelloagan & Martir, 2025; Park et al., 2026).

While traditional views center on what graduates should do, emerging scholarship advocates for a broader perspective that considers who graduates are becoming within their communities of practice (Becker, 2017; Haynes, 2026). This more holistic approach emphasizes the importance of values, dispositions, and professional thinking that accompany skills acquisition (Billettt, 2015; Tummons, 2014). Mezirow (1997) suggests adopting a more developmentally focused process of career readiness, which involves learners critically examining the assumptions that shape their professional experiences (Mezirow, 1997). This perspective shifts

the focus from merely acquiring competencies instead to fostering professional identity within authentic communities of practice (Cranton, 1997; Lave J. , 1991; Lave & Wenger, 1991). The hypothesis is that career readiness is cultivated through intentionally designed educational experiences that help learners interpret professional expectations and construct their identities, all of which enhance their integration into the workforce.

## Participation, Practice-Based Learning, and Learning Architectures

Lave and Wenger (1991) present learning as a process of legitimate participation where newcomers gradually become comfortable with robust involvement in professional communities (Lave & Wenger, 1991). This model contrasts with traditional classroom learning, positioning knowledge construction as a product of participation in activities, relationship building, and cultural experiences. Central to this view are practice-based and situated learning theories, which highlight the importance of experiential learning (Lave J. , 1991; Lave & Wenger, 1991). These strategies are essential for creative contexts like design studios, where students actively engage in skill acquisition (Cross, 1982; Faerm, 2023). This perspective shifts the focus from mere knowledge transmission to active participation, which is essential for developing competence and professional identity (Cranton, 1997; Mezirow, 1997; Shor & Freire, 1987).

Here, the concept of learning architectures is crucial, as it encompasses the various elements which influence how participation occurs in educational settings. Tummons (2014) describes learning architectures as a combination of instructional strategies, engaging communities, activities, resources, and expectations that create opportunities for students to engage authentically within their professional fields (Tummons, 2014). Effective design of these architectures requires educators to scaffold learning experiences that permit students to progress from foundational knowledge to more complex understandings integrated within real-world contexts.

Billett (2015) argues that practice-based learning is most effective when experiences are intentionally integrated with academic learning through preparation, guidance, participation, and reflection. This approach facilitates professional competence through engagement with authentic experiences, negotiation of meanings, allowing students to take on increasing responsibility in professional settings, all of which enhance both skills that allow for membership in professional communities.

Collectively, these perspectives suggest that meaningful participation, reflection, and feedback transform experience into learning and contribute to professional identity formation (Billett, 2015; Cesinger & Dittenberger, 2025; Chick, 2019; Grabove, 1997; Lave & Wenger, 1991; Tummons, 2014; van Halen-Faber, 1997).

## Dialogue, Critique, Reflection, and Transformative Learning

Billet (2015), Cranton (1997), and Mezirow (1997) caution that while participation can facilitate learning, it does not guarantee professional growth (Billett, 2015; Cranton, 1997; Mezirow, 1997). To foster meaningful development, learners must be able to interpret experiences, examine assumptions, negotiate meaning, and reflect upon their evolving understandings (van Halen-Faber, 1997; Wilcox, 1997). Dialogue plays a crucial role in this process, as articulated by Shor and Freire (1987) and Mezirow (1997). These authors argue

that dialogue encourages learners to challenge their assumptions and develop autonomous views of their professional development (Mezirow, 1997; Shor & Freire, 1987).

In studio-based disciplines like fashion design, critique serves as a structured form of dialogue that helps learners engage in professional norms, expectations, and disciplinary thinking. These steps are critical in supporting the transition from a fixed identity to that of a more confident and empowered professional (Adu et al., 2025; Cranton, 1997).

Reflection-in-action, proposed by Schön (1983), is another essential element in this transformational learning process (Schön, 1983; van Halen-Faber, 1997). Strategically including cycles of reflection allows students to convert their experiences into meaningful learning. Kolb (2015) indicates that reflection is critical component of experiential learning, while Cranton (1997) and van Halen-Faber (1997) note the significance of reflection as a key mechanism through which learners begin to question assumptions and developing new understandings of professional identities. (Cranton, 1997; Kolb, 2015; van Halen-Faber, 1997). Through critical self-reflection, learners begin to connect personal experience with broader professional meaning and values, fostering their growth within communities of practice (Lave & Wenger, 1991).

Transformative learning in these contexts occurs when individuals critically examine existing frames of reference and instead begin to adopt new perspectives (Mezirow, 1997). This transformation requires critical reflection and engagement in discourse, with autonomy, participation, and communication being essential components (Shor & Freire, 1987) (Grabove, 1997; Mezirow, 1997). Authentic experiences enable learners to link disciplinary knowledge with personal significance, while the uncertainty inherent in these processes can stimulate growth (Foster, 1997; Hodges, 2019). Collectively, the integration of dialogue, critique, reflection, and transformative learning highlights that professional development emerges from examining assumptions and negotiating meanings, all of which help to transform participation into meaningful learning and identity formation.

## Professional Identity Formation

Professional identity formation is a vital outcome of engaging in communities of practice, where learning extends beyond technical skill acquisition to include the development of one's professional self-perception. As learners navigate low-stakes experiences that mirror professional environments, they become socialized into the norms and expectations of their fields, reshaping their identities in relation to the professional community (Lave J. , 1991; Lave & Wenger, 1991; Mezirow, 1997). This process is further enriched by participation in diverse contexts, where learners cultivate appropriate vocabulary, internalize professional standards, and develop a sense of belonging, all of which facilitate the formation of their professional identities (Cranton, 1997; Hodges, 2019; Mezirow, 1997; Wenger-Trayner et al., 2014).

Transformative learning further highlights the importance of identity development, emphasizing how critical reflection and engagement with contrasting viewpoints can lead learners to revise their assumptions and foster more autonomous professional identities (Cranton, 1997; Grabove, 1997; van Halen-Faber, 1997). Billet (2015) argues that practice-based learning fosters not only competence but also confidence, and professional agency (Billett, 2015). As learners assume greater responsibility for their decisions and actions, they begin to view themselves as legitimate contributors to their fields, enhancing their participation in professional practice (Chick, 2019).

When taken as a whole, communities of practice, practice-based learning, and transformative learning suggest that professional identity formation serves as a critical bridge between educational experiences and career readiness outcomes (Wenger-Trayner et al., 2014). In this context, participation, dialogue, critique, reflection, and transformative learning operate as interconnected processes through which learners begin to think, act, and participate more fully as members of their chosen professions (Billett, 2015; Cranton, 1997; Mezirow, 1997).

### Toward a Curriculum Design Framework for Career Readiness

Collectively, this body of literature suggests that career readiness emerges through a developmental process that extends beyond technical skills acquisition. Communities of practice, practice-based learning, transformative learning, and learning architecture perspectives each emphasize different dimensions of professional development, yet together point towards an emphasis on participation, reflection, and identity formation as central mechanisms through which learners prepare for professional practice (Billett, 2015; Cranton, 1997; Foster, 1997; Gibson & Mader, 2019; Grabove, 1997; Haynes, 2026; Lave & Wenger, 1991; Mezirow, 1997; Shor & Freire, 1987; Tummons, 2014; Wenger-Trayner et al., 2014).

The literature included above suggests that there is a sequential relationship across these varied forms of educational processes. Intentional curriculum design promotes opportunities for participation, while participation provides access to professional communities and authentic practice. Additionally, dialogue, critique, and reflection transform participation into deeper learning; while collectively these experiences contribute to professional identity formation in support of more effective career readiness.

From this perspective, career readiness is less a collection of isolated skills and competencies, but instead an emergent outcome of intentionally designed learning environments which include participation, reflection, professional socialization, and identity development. This strategy lays the groundwork for a theoretical foundation for curriculum design, which is explored as the framework used in this study.

## Methodology

### Research Design

This study employed a qualitative case study design to examine how an industry- embedded capstone experience contributed to professional identity formation and career readiness among upper-level fashion design students. Based on the work of Groat and Wang (2013), a case study methodology was selected for this research because of its flexible format for capturing insights into the phenomena that takes place within specific situations (Groat & Wang, 2013).

The study focused on a single capstone model implemented across two sequentially linked fashion design courses, and examined how students experienced participation, critique, reflection, and recurring cycles of industry engagement throughout the academic year.

The study used this capstone model as a case through which to investigate how scaffolded participation, critiques, reflection, and authentic engagement with industry professionals contribute to professional identity formation and career readiness outcomes.

## Educational Context

The study took place at a public institution of higher learning located within the southern part of the United States. Situated within that organization's college of art and design, the study focuses primarily on the experiences of upper-level fashion design students (n:40) who were enrolled in their final year capstone courses with the goal of earning a BFA in fashion design. The courses which were involved took place over both the fall and the spring terms, as part of the fashion design program's sequentially oriented degree plan.

The capstone sequence serves as the culminating educational experience within the program and requires students to synthesize prior learning through the development of a professional apparel design collection and participate in a formal presentation process. Students enrolled in this program often do so with the intention of entering the fashion industry post-graduation, preferably within the design industries that focus on apparel product development.

## Industry-Embedded Learning Architecture

The instructional design for this study included a learning architecture which included strategically scaffolded milestones of industry engagement. At the beginning of the academic term, students participated in a low-stakes Career Expo, where they interacted informally with regional employers and received feedback on their portfolios and resumes. This was followed by a formal critique of their conceptual development ideas from an invited industry professional. In preparation, the instructor equipped students with training that prompted them to anticipate potential questions or challenges in their presentations, which provided early exposure to the expectations and standards common to the fashion industry.

Throughout the term, the instructor implemented several low-stakes checkpoints, fostering continuous feedback as the students refined their garment prototypes and presentation materials. The midpoint of the term included another round of industry critique with different professionals to challenge students' assumptions and encourage diverse viewpoints. The academic term concluded with a formal defense of the students' thesis collections in a format resembling an industry trade show, allowing the students to experience authentic professional engagement. This carefully sequenced progression aimed to facilitate participation, dialogue, personal reflection, and professional socialization, aligned within the broader learning architecture aimed at achieving professional identity formation.

## Data Sources

Throughout the study, several forms of data were collected to build a cohesive understanding of the phenomena unfolding throughout the course experience. From the students, both pre-survey (n:40) and post-survey (n:38) data was collected. The findings which emerged from these surveys was then blended with personal narratives which the students provided based on instructor developed reflection prompts. Together, these findings were blended to gain a more accurate understanding of the student perspective.

Meanwhile, the lead instructor took field notes which included observations of the students inside the classroom, at the career expo, and during critique settings. Photo documentation supplemented these findings, which was correlated back to the observational findings and the field notes.

Additional information was collected by external reviewers involved in the study, which included the industry expert guest critics (n:2), career expo participants (n:12), and the industry jury panelists who reviewed the student's thesis collections (n:20). These data sets included survey data, along with written evaluations, and transcripts from the critique sessions.

Using a mixed methods approach that blended the qualitative data which was collected from student surveys, student reflections, instructor observations, and feedback from external industry professionals, information was analyzed for emerging themes and coded using keyword analysis. The findings were then cross-examined by another researcher within the institutions' design department to verify the analysis via consensus.

### Data Analysis

Based on the work of Boyatzis (1998), the qualitative data that was collected was then analyzed using an inductive thematic analysis approach (Boyatzis, 1998). Student reflections, survey comments, instructor observations, and industry feedback data sources where then reviewed with the goal of finding emerging themes or concepts related to professional development.

Initial codes were grouped into broader thematic categories which aligned with the findings from the literature review, such as communication, critical thinking, professional judgement, confidence, reflective practice, and agency. These themes were then compared across data sets to identify possible areas of agreement, convergence of themes, and divergent areas or flashpoints hidden within the data.

### Limitations

While the findings largely aligned with the information which emerged from the literature review, there are several limitations which may have impacted the reliability of the findings from this study, the greatest of which may be researcher bias. Because the lead instructor also served as the primary researcher, interpretations of the findings may have been influenced by proximity to the educational context. This issue also applies to the use of self-reported data, either from the instructor's observations or the student's self-reflections.

Another issue emerges from the limited transferability of the findings, since the study took place only within a single institutional setting, potential application across other disciplines may be limited.

## Findings

Once the data were coded, seven key themes emerged from the findings. Four overarching constructs emerged from the themes: Cluster A: Professional Communication and Socialization. Cluster B: Judgement, Reflection, and Transformative Learning. Cluster C: Confidence and Agency. Cluster D: Professional Identity Formation.

## Cluster A: Professional Communication and Socialization

### *Theme 1: Professional Communication*

Student responses frequently referenced networking opportunities, industry panels, and professional interactions as valuable aspects of the program experience, such as "working the fashion shows and industry panels."

Many respondents also reported high levels of confidence in presentation and communication skills, suggesting that repeated exposure to professional audiences supported communication development, stating "(the program) does a great job at networking and getting connections for their students."

Over time, the students became more at ease with interacting in professional settings, as they strengthened their vocabulary and learned to align their materials to industry expectations. These incrementally staged milestones functioned as "dress rehearsals" before students arrived at their final thesis defense. With repetition, students honed their defense "scripts", and the more formal modes of presentation and critique became more natural. Students indicated increased familiarity with appropriate industry terms and mannerisms over time.

### *Theme 2: Professional Community and Belonging*

Several students repeatedly referenced terms such as: "classmates", "collaboration", "networking", and "industry connections." For example, several students cited: "working with my classmates," and "making connections." These findings align with the concept of communities of practice. The instructor also noted several groups of students pooled their resources and formed workgroups that worked as accountability partners in meeting project deadlines or advancing the quality of execution in their work.

The findings here are also corroborated with the industry critic feedback from the jury panel evaluations. Across the critique evaluations, reviewers repeatedly referenced their observations of the students as "very professional and well spoken", and able to "articulate ideas and process very professionally." These findings may also suggest indicators of career readiness, including mentions in their assessments of the students that they felt they were "ready to work in the industry," and an "excellent addition to a design team."

## Cluster B: Judgement, Reflection, and Transformative Learning

### *Theme 3: Critical Thinking and Professional Judgement*

Here, students indicated that they had begun to move beyond technical execution to develop their own process of problem solving according to professional criteria. Several students reported that they experienced "stumbling blocks" or moments of frustration early in the process when their attempts to execute a particular technique failed to meet expectations. Terminology such as "troubleshooting", "fit challenges", "technical problem solving", and "adapting designs" were used repeatedly.

Additionally, many students reported a strong desire for more opportunities to experience "fitting, technical design, grading, and production techniques." These comments imply that

students beginning to move beyond simple garment construction challenges and instead recognize the complexity involved in professional decision-making.

### *Theme 4: Reflection and Self-Awareness*

Students became increasingly aware of their strengths, weaknesses, and professional goals, citing that although they felt they had achieved mastery enough to successfully complete the degree plan, there were several skills they wish they had yet learned. The instructor also observed that as the capstone progressed, some students began to hyper-focus on their personal shortcomings considering frustrations while attempting to meet the project development timeline.

However, rather than viewing these comments as complaints, they can be interpreted as indicators of increased professional awareness. These emerging indicators of challenge and frustration also showcase that students were beginning to recognize the misalignment between what knowledge and skills they had mastered and compare that with what they still needed to learn.

## Cluster C: Confidence and Agency

### *Theme 5: Confidence Through Repeated Participation*

Low-stakes engagement gradually reduced anxiety and increased confidence. Going into the final preparations before their collection defense jury panel event, several students indicated that they "felt ready", that they were "feeling prepared for industry roles," suggesting that they felt increasingly prepared to meet industry expectations. One student stated: "I learned so much not only about myself but also about the design process."

The findings here are also corroborated with the industry critic feedback from the jury panel evaluations. Across the critique evaluations, reviewers repeatedly referenced their observations of the student's "confidence", "professionalism", and their "ability to articulate ideas". Some critics stated that "the designer came across as confident and well prepared."

### *Theme 6: Agency and Professional Responsibility*

Students increasingly assumed ownership of their decisions and outcomes. By the end of the academic term, several of the students reported how they could now "see themselves in the industry." Several students took the initiative to begin exploring post-graduate career opportunities after the first six months, refining their resumes and embarking on the job interview process. In their exit survey, numerous students reported that they had enjoyed active participation in internships, employment opportunities, and professional interviews, suggesting increased ownership of career development and professional decision-making.

## Cluster D: Professional Identity Formation

### *Theme 7: Emerging Professional Identity*

Several students moved beyond discussing technical competencies and instead described evolving understandings of who they were becoming designers. These comments strongly suggest professional identity formation, such as in the student statements where they mentioned

that they were "learning to hone my personal aesthetic," or that "I learned so much not only about myself."

One student also highlighted that they felt that the learning architecture had enabled and empowered them by "Being in an environment that encouraged us to be very experimental and show off our own style (aesthetic)."

## Discussion

Career readiness is not the outcome of isolated skills training but emerges through intentionally designed learning architectures that support participation, reflection, professional socialization, agency, and professional identity formation.

### Participation in Professional Socialization

This cluster emerges from the findings which mention a combination of networking, industry interaction, repeated cycles for the demonstration of presentation skills, and external validation from industry professionals as critics. These findings suggest that repeated engagement with industry professionals can function as a process of professional socialization through which students become increasingly comfortable participating in disciplinary specific communities of practice (Wenger, 1998). These findings support Lave and Wenger's (1991) proposition that learning occurs through participation rather than knowledge acquisition alone, and Wenger-Trayner et al.'s (2014) argument that learning takes place within the broader landscapes of practice (Lave & Wenger, 1991; Wenger-Trayner et al., 2014). Rather than functioning as isolated networking activities, these encounters appear to have functioned as legitimate opportunities for authentic participation through which students gradually transitioned from observers of professional expectations to instead become increasingly confident participants when interacting within professional communities. (Lave & Wenger, 1991).

### Critique, Reflection, and Transformative Learning

This cluster emerges from findings which mention terms like "trouble shooting", challenges, adaptive learning, and recognizing skills gaps, highlighting that participation alone may not be sufficient (Cranton, 1997). These findings are consistent with the work of Mezirow's (1997) Transformative Learning, whereby instead of participation alone, professional growth emerges through cycles of critique, reflection, and moments of frustration which challenge existing underlying assumptions (Mezirow, 1997). Instead, student's awareness of skill limitations suggest that uncertainty may have helped function as a catalyst for learning rather than merely function as an obstacle (Hodges, 2019).

These insights are best gleaned through iterative cycles of reflection and action, whereby reflection helps promote transformation in one's perspectives when coupled with critique, reflection, and the challenge of navigating moments of uncertainty (Billett, 2015; Cranton, 1997; Grabove, 1997; Kolb, 2015; Schön, 1983; van Halen-Faber, 1997).

### Confidence, Agency, and Professional Becoming

The findings suggest that confidence and agency developed through repeated engagement with authentic professional experiences. As students moved through increasingly demanding stages of participation, they became more willing to assume responsibility for decision-making, seek

professional opportunities, and independently navigate career pathways. These findings closely align with Billett's (2015) argument that practice-based learning contributes not only to competence development, but also confidence, judgement, occupational understanding, and agency (Billett, 2015).

Notably, however, confidence and agency did not appear simultaneously within the findings. Instead, students moved through a progression whereby they first began expressing increased confidence when communicating with industry professionals when defending their decisions. Over time, confidence became an attribute which supported the growth of their personal agency, which was reflected through their job seeking behaviors, participation in interviews, and independent career planning. This progression suggests that confidence may serve as a developmental precursor which then evolves into professional agency.

**Professional Identity Formation as the Bridge to Career Readiness**

One significant finding from this study is the role of professional identity formation as a bridge between educational experience and career readiness. Students increasingly identified themselves as "designers", suggesting that career readiness can be understood as a process of becoming a professional rather than merely acquiring skills.

This shift from competence acquisition to professional self-identification aligns with Wenger's (1998) assertion that learning involves becoming a participant within a professional community (Wenger, 1998). Students reported not just enhanced competence but also a growing sense of legitimacy within the profession.

Professional identity formation appears to be the mechanism through which educational experiences translate into career readiness outcomes. Ultimately, career readiness emerges through identity development fostered by intentional cycles of active participation, reflection, critique, and professional socialization (Kolb, 2015; (Schön, 1983; Wilcox, 1997).

**Toward a Curriculum Framework for Career Readiness**

These findings suggest that career readiness is not primarily the result of skills acquisition but instead emerges through intentionally designed learning architectures that support participation, professional identity formation, and becoming a professional. Based on the findings and supporting literature, the study proposes a curriculum design framework in which career readiness emerges through a developmental sequence of participation, critique, reflection, transformation, and identity formation.

The proposed framework illustrates a strategically scaffolded developmental sequence through which educational experiences contribute to career readiness outcomes (see Figure 1, included below). Within this framework, learning architectures provide opportunities for participation. These opportunities for participation introduce learners to professional communities and authentic behaviors in professional practice settings. Critique, dialogue, and reflection combine to help transform the participation into self-directed learning, while repeated engagement supports the development of confidence and agency. Together, these experiences contribute to professional identity formation, which functions as the bridge between educational experiences and career readiness.

Although this case study emerged from fashion design education, the framework may have broader significance across disciplines that seek to prepare students for professional practice through authentic engagement, reflection, and participation in professional communities. While the specific learning activities may vary across disciplines, the underlying sequence of participation, reflection, transformation, and identity formation may provide a transferable framework for curriculum design in a wide range of professional contexts.

**Figure 1**
*Professional Identity Formation as a Bridge to Career Readiness*

*Source*: Trippeer, Barbara (2026); *Note*. Conceptual model showing the developmental relationship between educational experiences, professional identity formation, and career readiness outcomes.

## Conclusion

The findings suggest that career readiness emerges through intentionally designed learning architectures that foster professional participation, socialization, reflection, and identity formation. Through scaffolded experiences that incorporated critique, reflection, and authentic professional engagement, students developed increasing confidence, agency, and professional identity. These findings support the proposed curriculum framework and suggest that educators across disciplines may benefit from designing learning architectures that prioritize authentic participation and professional socialization as pathways toward career readiness through professional identity formation.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

Generative artificial intelligence tools were used in the preparation of this manuscript to assist with spelling, copy editing, and stylistic grammatical refinement of author-generated content. AI was also used to assist in refining the final execution of the visual image used for the Figure Model. All conceptual development, literature selection, analysis, interpretation of findings, and final manuscript preparation were conducted by the author. The author reviewed and verified all content, and assumes full responsibility for the accuracy, integrity, and originality of the work presented here.

## References

Adu, C., Biney-Aidoo, V., Okai-Mensah, C., & Adu Akwaboah, D. (2025). Competency-Based Training in Fashion Design Education: Graduate Competencies and Job Performance. *International Conference on Applied Arts, Design, and Sustainability (ICAADS 2025).* Springer Nature. https://doi.org/10.2991/978-2-38476-527-0_5

Becker, B. J. (2017). Chapter 10: Improving the Design and Use of Meta-Analyses of Career Interventions. In J. Sampson, E. Bullock-Yowell, V. Dozier, D. Osborn, & J. Lenz, *Integrating theory, research, and practice in vocational psychology: Current status and future directions* (pp. 95–107). Florida State Open Publishing. https://doi.org/10.17125/svp2016.ch10

Billett, S. (2015). *Integrating Practice-based Experiences into Higher Education.* Springer Dordrecht. https://doi.org/10.1007/978-94-017-7230-3

Boyatzis, R. E. (1998). *Transforming qualitative information: thematic analysis and code development.* Sage Publications.

Cesinger, B., & Dittenberger, S. (2025). Shaping Entrepreneurial Design Identity in Design Education. *She Ji: The Journal of Design, Economics, and Innovation, 11*(1), 61–75. https://doi.org/10.1016/j.sheji.2025.01.003

Chick, N. L. (2019). The Crossroads of SoTL and Signature Pedagogies. In C. Gibson, & S. Mader, *Building Teaching and Learning Communities: Creating Shared Meaning and Purpose* (pp. 41–50). Association of College & Research Libraries. ISBN 978-0-8389-4656-5

Cranton, P. (Ed.). (1997). *Transformative Learning in Action: Insights from Practice* (Vol. 74). Jossey-Bass Publishers. ISBN-10: 0787998419

Cross, N. (1982). Designerly ways of knowing. *Design Studies, 3*(4), 221–227. https://doi.org/10.1016/0142-694X(82)90040-0

Cross, N. (2007). Understanding Design Cognition. In *Designerly Ways of Knowing* (pp. 77–93). Springer.

Faerm, S. (2023). *Introduction to Design Education: Theory, Research, and Practical Applications for Educators.* Routledge. https://doi.org/10.4324/9781003049166

Foster, E. (1997). Transformative Learning in Adult Second Language Learning . In P. Cranton, *Transformative Learning: Insights from Practice* (pp. 33–40). Jossey-Bass Publishers. ISBN-10: 0787998419

Gelloagan, A. E., & Martir, E. M. (2025, January). Students' Industry Preparedness, Job Proficiency, and Academic Achievement. *International Journal of Science and Management Studies, 8*(1), 39–47. https://doi.org/10.51386/25815946/ijsms-v8i1p105

Gibson, C., & Mader, S. (2019). *Building Teaching and Learning Communities: Creating Shared Meaning and Purpose.* Association of College & Research Libraries. ISBN 978-0-8389-4656-5

Grabove, V. (1997). The Many Facets of Transformative Learning Theory and Practice. In P. Cranton, *Transformative Learning in Action: Insights from Practice* (pp. 89–102). Jossey-Bass Publishers. ISBN-10: 0787998419

Groat, L. N., & Wang, D. (2013). Chapter 12: Case Studies and Combined Strategies. In L. N. Groat, & D. Wang, *Architectural Research Methods* (Second ed., pp. 415–468). Wiley. ISBN: 978-0-470-90855-6

Haynes, D. (2026, July 31). *Helping Students Find Connection and Purpose, Not Just Careers.* Retrieved August 05, 2026, from HigherEdJobs: https://www.higheredjobs.com/blog/postDisplay.cfm?post=4783&blog=49&Title=Helping%20Students%20Find%20Connection%20and%20Purpose%2C%20Not%20Just%20Careers

Hodges, L. C. (2019). Embracing the Liminality of Pedagogical Change through the Scholarship of Teaching and Learning. In C. Gibson, & S. Mader, *Building Teaching and Learning Communities: Creating Shared Meaning and Purpose* (pp. 27–40). Association of College & Research Libraries. ISBN 978-0-8389-4656-5

Kolb, D. A. (2015). *Experiential learning: experience as the source of learning and development* (Second ed.). Pearson Education, Inc.

Lave, J. (1991). Situated Learning in Communities of Practice. In L. B. Resnick, J. M. Levine, & S. D. Teasley, *Socially Shared Cognition* (1st ed., pp. 63–82). American Psychologocial Association.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press.

Mezirow, J. (1997). Transformative Learning: Theory to Practice. In P. Cranton, *Transformative Learning in Action: Insights from Practice* (pp. 5–12). Jossey-Bass Publishers. ISBN-10: 0787998419

Park, J. G., Gam, H. J., Trippeer, B. C., & Stark, J. (2026, June 2). Investigations regarding perspectives between design education and industry practices in sustainability design. *International Journal of Sustainability in Higher Education, 27*(6), 1310–1328. https://doi.org/10.1108/IJSHE-07-2024-0490

Schön, D. A. (1983). *The Reflective Practitcioner: How Professionals Think in Action.* Routledge.

Shor, I., & Freire, P. (1987, October). What is the "dialogical method" of teaching? *Journal of Education, 169*(3), 11–31. https://doi.org/10.1177/002205748716900303

Tummons, J. (2014). Learning architectures and communities of practice in higher education. In *Theory and Method in Higher Education Research II* (Vol. 10, pp. 121–139). Emerald Group Publishing Limited. ISBN 9781350130975

van Halen-Faber, C. (1997). Encouraging Critical Reflection in Preservice Teacher Education: A Narrative of a Personal Learning Journey. In P. Cranton, *Transformative Learning in Action: Insights from Practice* (pp. 51–60). Jossey-Bass Publishers.https://doi.org/10.1002/ace.7406

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity.* Cambridge University Press.https://psycnet.apa.org/doi/10.1017/CBO9780511803932

Wenger-Trayner, E., Fenton-O'Creevy, M., Hutchinson, S., Kubiak, C., & Wenger-Trayner, B. (2014). *Learning in landscapes of practice: Boundaries, identity, and knowledgeability in practice-based learning.* Routledge. https://doi.org/10.4324/9781315777122

Wilcox, S. (1997). Becoming a Faculty Developer. In P. Cranton, *Transformative Learning in Action: Insights from Practice* (pp. 23–31). Jossey-Bass Publishers. ISBN-10: 0787998419

# Non-cognitive Factors of Mathematics Achievement Among Higher Education Students in Bahrain

James Ray Mendaje, Bahrain Polytechnic, Bahrain

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This study examined the non-cognitive factors influencing mathematics achievement among higher education students in Bahrain through a descriptive-correlational design, employing correlation analysis and structural equation modeling (SEM) to test the predictive roles of mathematics anxiety, attitude towards mathematics, motivation, and learning strategies. Results indicated that students generally exhibited moderate levels of mathematics anxiety, attitude, and motivation, alongside moderately effective use of learning strategies. Correlation analysis revealed that mathematics anxiety was negatively correlated to attitude, motivation and achievement, and positively linked to learning strategies. In contrast, attitude, motivation, and learning strategies were positively associated with one another and with achievement. The best-fit SEM confirmed mathematics anxiety, attitude, and learning strategies as significant predictors of achievement. These findings underscore the importance of strengthening positive attitudes, enhancing learning strategies, and reducing anxiety to improve mathematics performance.

*Keywords:* non-cognitive factors**,** mathematics achievement, best-fit SEM, correlation analysis, higher education students

iafor
The International Academic Forum
www.iafor.org

## Introduction

Mathematics is essential for solving real-world problems and supports advancement in fields like science, engineering, economics, and technology (Sharma, 2021). It develops critical skills such as problem-solving, critical thinking, and communication, helping students make informed decisions (Khan & Salman, 2020; Sharma, 2021; Wanamaker, 2024). Despite its importance, achieving proficiency in mathematics remains a global challenge, with recent reports emphasizing the need for educational reforms to improve math achievement and support both individual and societal progress (Seagull, 2024).

In the 2023 Trends in International Mathematics and Science Study (TIMSS), several countries consistently outperformed others in both fourth and eighth grade mathematics, with many exceeding the international average, while some regions performed below it. Within the Gulf Cooperation Council, Bahrain ranked mid-range for both grades but remained below the international average, underscoring the need for continued educational improvements (von Davier et al., 2024).

Bahrain's performance in the TIMSS from 2003 to 2023 shows overall improvement in mathematics achievement up to 2019 at both fourth- and eighth-grade levels, followed by declines through 2023, with some earlier fluctuations (von Davier et al., 2024). Regionally, Bahrain ranked among the top Gulf Cooperation Council (GCC) countries, although all GCC nations remained below the international average. Despite progress and ongoing government reforms in curriculum, teacher training, and policy, Bahrain's mathematics achievement still falls short of global standards, highlighting the need for continued improvement.

Mathematics anxiety remains a significant issue in education, negatively affecting students' learning experiences, well-being, and long-term engagement with mathematics, with evidence showing a notable increase in anxiety levels across OECD countries between 2012 and 2022 (Earp, 2024; OECD, 2024). Students' attitudes and confidence toward mathematics are equally important, as low confidence can limit lifelong learning and reduce interest in math-related careers (OECD, 2024). Motivation further influences achievement, as confident students tend to be more intrinsically motivated and use effective learning strategies, while less confident students rely more on extrinsic factors, though many still recognize the value of mathematics (OECD, 2024). In the context of Bahrain educational landscape, mathematics anxiety, attitude towards mathematics, and learning strategies remain an underexplored area of study. No research has been conducted to explore students' motivation in connection to learning mathematics.

Despite extensive research on mathematics anxiety, attitudes, motivation, and learning strategies, studies often examine these factors in isolation, leaving a gap in understanding their combined and interactive effects on academic achievement. This gap is more evident in the Gulf Cooperation Council (GCC) region, particularly in Bahrain, where limited and outdated research has explored these non-cognitive factors among higher education students. Moreover, the lack of advanced statistical techniques in existing studies further restricts the depth and reliability of findings, highlighting the need for more comprehensive and methodologically robust research.

This study is anchored on theories explaining key non-cognitive factors influencing mathematics achievement. Mathematics anxiety is framed by the deficit theory and the debilitating anxiety model, which explain how poor performance and cognitive interference

contribute to anxiety (Brooks, n.d.). Attitudes toward mathematics are grounded in Ajzen's Theory of Planned Behavior, emphasizing affective, cognitive, and behavioral dimensions that influence students' engagement with mathematics (Ajzen, 1991; Fennema & Sherman, 1978; Mohr-Schroeder et al., 2017; Sengül & Dereli, 2013). Motivation is supported by expectancy-value theory, self-determination theory, and self-efficacy theory, highlighting the roles of expectations, intrinsic and extrinsic factors, and beliefs in one's abilities (Bandura, 1985; Eccles et al., 1993; Ryan & Deci, 2000). Learning strategies are supported by metacognitive, social learning, and experiential learning theories, which emphasize cognitive regulation, observation, and experience-based learning (Bandura, 1985; Kolb, 1984). The overall framework of the study is further substantiated by Ajzen's Theory of Planned Behavior (TPB), which provides a foundation for understanding how students' attitudes, motivation, anxiety, and learning strategies influence mathematics achievement.

The study addressed the following questions: What is the extent of mathematics achievement of higher education students in Bahrain in terms of final course grades; what is the level of non-cognitive factors in terms of (a) mathematics anxiety, (b) attitude towards mathematics, (c) motivation, and (d) learning strategies; is there a significant relationship between and among mathematics achievement and these non-cognitive factors; and what is the best-fit structural equation model for mathematics achievement as influenced by mathematics anxiety, attitude towards mathematics, motivation, and learning strategies?

## Literature Review

### Mathematics Achievement

Academic achievement is a widely researched topic in education, psychology, and social sciences, as it serves as a key indicator of student success and future opportunities. According to Moss & Moss-Racusin (2021), it refers to an individual's scholastic accomplishment, distinct from innate intelligence. Mathematics achievement is a construct used to represent the level or nature of demonstrated learning in relation to a specified domain of mathematics. It is typically evaluated using standardized tests or structured assessment methods (Shear, 2020). Numerous studies have been conducted to examine the factors influencing mathematics achievement. Villa and Sebastian (2021) demonstrated that achievement motivation plays a crucial role in predicting students' performance in mathematics. Complementing this, Abin et al. (2020) highlighted cognitive, motivational, and emotional components as significant determinants of mathematics achievement, showing that a model integrating these factors accounted for considerable differences in students' math performance. Additionally, previous academic performance in mathematics, as well as critical and creative thinking skills, have been identified as predictors of math achievement (Fernández-Cézar et al., 2021). Moreover, Mendaje (2018) found that students' level of learning engagement and sustained participation in mathematics learning were significantly and positively associated with mathematics achievement, suggesting that learning-related factors contribute to improved mathematical performance.

### Mathematics Anxiety

Mathematics anxiety is the feeling of tension, fear, and apprehension experienced when engaging in mathematical tasks, which can impair performance and lead to avoidance of mathematics-related activities (Ashcraft, 2002; Hembree, 1990; Richardson & Suinn, 1972). It may occur as a temporary state during mathematical tasks or as a persistent trait that

influences long-term attitudes and behaviors toward mathematics (Buckley et al., 2016). As a multidimensional construct involving emotional, cognitive, physical, and behavioral responses, mathematics anxiety negatively affects students' attitudes, learning, and achievement in mathematics (Caviola et al., 2019; Sangaram & Maat, 2020; Telaumbanua et al., 2023).

### Attitude Towards Mathematics

Attitude toward mathematics is a multidimensional construct that encompasses students' cognitive, emotional, and behavioral dispositions toward the subject (Wen & Dubé, 2022). It includes key components such as self-confidence, perceived value, and enjoyment, which influence students' engagement and achievement in mathematics (Tapia & Marsh, 2002). Students with greater confidence, stronger perceptions of the value of mathematics, and higher enjoyment are more likely to participate actively in learning, persist in mathematical tasks, and achieve better academic outcomes (OECD, 2024).

### Motivation

Motivation refers to the psychological processes that initiate, direct, and sustain goal-oriented behavior, influencing students' effort, persistence, and academic performance (Lens & Vansteenkiste, 2020; Schunk & DiBenedetto, 2020). It may be intrinsic or extrinsic and is particularly important in mathematics, where perseverance and self-discipline are essential for success (Kumar, 2020). Research by Davidovitch and Dorot (2023) further showed that motivation is positively associated with age and academic performance, with undergraduate students demonstrating higher motivation than high school students.

### Learning Strategies

Learning strategies are cognitive processes and behaviors employed by learners to enhance their academic achievement. They can be classified into basic strategies used by younger children and advanced strategies developed by older students, involving metacognition (Yip, 2019). However, students' strategy use varies depending on factors such as assessment activities, and those with low academic performance often struggle with organization and strategy knowledge (García-Pérez et al., 2020). Importantly, the ability to transfer deep learning strategies between courses is associated with higher grades, while continued use of surface strategies may lead to lower achievement. These findings emphasize the need for educators to support students in developing and transferring effective learning strategies, particularly in problem-based learning environments (Saqr et al., 2023).

### Relationship Between Mathematics Anxiety and Mathematics Achievement

The reviewed studies consistently show that mathematics anxiety is negatively associated with mathematics achievement, whereas positive attitudes toward mathematics are generally linked to better academic performance. Meta-analyses and empirical studies (e.g., Baral et al., 2025; Caviola et al., 2022; Zhang et al., 2019) found that higher mathematics anxiety leads to lower achievement, although motivation, working memory, age, and other cognitive and psychological factors also influence performance. Similarly, studies by Dowker and Sheridan (2022), Hwang and Son (2021), and Kiwanuka et al. (2020) reported that positive attitudes toward mathematics support higher achievement, while Libradilla et al. (2023) found no significant relationship, suggesting that the effect of attitudes may vary across contexts.

Overall, the evidence highlights the importance of reducing mathematics anxiety and fostering positive attitudes to enhance students' mathematics achievement.

**Relationship Between Motivation and Mathematics Achievement**

Several studies have identified motivation as an important factor influencing mathematics achievement. Students with higher intrinsic and extrinsic motivation, self-efficacy, goal orientation, and educational aspirations demonstrated greater engagement and better academic performance (Akayuure & Akayuure, 2024; Asanre, 2024; Jehadus et al., 2022; Ovella, 2023; Shum & Fryer, 2023). Similarly, Bakar et al. (2022) and Oluyomi et al. (2024) confirmed the positive influence of motivation on achievement, highlighting its important role in enhancing students' success in mathematics.

**Relationship Between Learning Strategies and Mathematics Achievement**

Research on mathematics education consistently emphasizes the importance of learning strategies in enhancing students' academic performance. Effective learning strategies, including cognitive, self-regulated, metacognitive, and effort management strategies, were consistently associated with higher mathematics achievement and improved academic performance (Almoslamani, 2022; Ha et al., 2023; Liebendörfer et al., 2022; Okudan & Yeşilyurt, 2023; Wild & Neef, 2023). These findings highlight the importance of fostering effective learning strategies to enhance students' success in mathematics.

**Relationship Between Mathematics Anxiety and Attitude Towards Mathematics**

Research consistently highlights the close relationship between mathematics anxiety and attitudes toward mathematics. Caputol and Chan (2024) found that attitudes toward mathematics significantly influence mathematics anxiety, emphasizing the role of students' emotional and perceptual experiences in shaping anxiety. Similarly, Alias et al. (2023) and Tuncer and Yilmaz (2020) reported a significant negative relationship between mathematics anxiety and attitudes, suggesting that reducing anxiety may foster more positive attitudes and improve academic achievement.

**Relationship Between Mathematics Anxiety and Motivation**

The relationship between mathematics anxiety and motivation has been widely examined in mathematics education research. Schukajlow et al. (2023) found that higher motivation helps reduce the negative effects of mathematics anxiety on learning and performance. Similarly, Süren and Kandemir (2020) and Supriadi et al. (2024) reported significant relationships between mathematics anxiety and motivation, indicating that stronger motivation may help mitigate anxiety and support better mathematical learning.

**Relationship Between Mathematics Anxiety and Learning Strategies**

Several studies have examined the relationship between mathematics anxiety and learning strategies, with mixed findings. Caputol and Marticion (2021) found significant positive relationships between mathematics anxiety and learning strategies, particularly motivated learning, self-regulation, and effort regulation. In contrast, Gursel et al. (2023) reported no significant relationship between mathematics anxiety and students' self-

regulated learning strategies, suggesting that this relationship may vary across educational contexts.

### Relationship Between Attitude Towards Mathematics and Motivation

Ampadu and Anokye-Poku (2022) investigated the influence of personal, motivational, and learning environment factors on students' attitudes toward mathematics among junior secondary school students. The study found that students generally had positive attitudes toward mathematics, which were positively associated with academic achievement and supported by motivation, self-efficacy, and a positive learning environment.

### Relationship Between Attitude Towards Mathematics and Learning Strategies

The association between attitudes toward mathematics and learning strategies has received considerable attention in recent research. Siller and Ahmad (2024) found that collaborative learning strategies improved both mathematics achievement and students' attitudes toward mathematics. Similarly, Caputol and Chan (2024) and Pizon and Ytoc (2021) reported significant positive relationships between learning strategies and attitudes, suggesting that effective learning approaches contribute to more positive perceptions of mathematics.

### Relationship Between Motivation and Learning Strategies

Molnár and Kocsis (2023) studied academic achievement in higher education by examining cognitive and non-cognitive factors. The study found that learning motivation and learning strategies positively predicted academic achievement and were positively correlated with each other, while learning attitudes were also positively associated with learning motivation.

## Methodology

### Research Design

This study utilized quantitative, non-experimental research design, specifically descriptive-correlational methodology to investigate the factors that would influence the mathematics achievement of higher education students in Bahrain. Strunk and Mwavita (2024) note that quantitative research involves the use of structured tools, such as surveys and experiments, to collect data that can be quantified. Bryman (2016) explained that quantitative design is characterized by its emphasis on measurement, statistical analysis, and the use of large samples to ensure the generalizability of findings. The descriptive-correlation design is a quantitative research approach that is used to explore relationships or associations between two or more variables without establishing causality (Salkind, 2017). This approach does not infer causation but rather identifies the degree and direction of relationships. This is essential for generating insights that could guide further research or practical applications (Creswell & Creswell, 2018).

The descriptive components of this study aimed to provide a comprehensive analysis of the respondents' levels or extent in five key areas: mathematics anxiety, motivation, attitude towards mathematics, learning strategies, and mathematics achievement. Specifically, it examined mathematics achievement in terms of mathematics course final grades, mathematics anxiety in terms of learning math anxiety and math evaluation anxiety, and attitudes towards mathematics in terms of self-confidence, value, and enjoyment.

Furthermore, the descriptive component evaluated motivation through dimensions such as intrinsic and extrinsic goal orientation, task value, control of learning beliefs, and self-efficacy for learning and performance. Additionally, it investigated the extent to which learning strategies were employed, encompassing rehearsal, elaboration, organization, critical thinking, metacognitive self-regulation, time and study environment, effort regulation, peer learning, and help seeking. The correlational aspect of the study aimed to explore how mathematics anxiety, mathematical attitudes, motivation, and learning strategies are interconnected.

Bahrain Polytechnic was selected as the primary locale for this study within Bahrain's higher education sector, which consists of 16 institutions, including 2 public (12.5%) and 14 private (87.5%) universities, with the Polytechnic representing 6.25% of the total. The study utilized data from students enrolled at Bahrain Polytechnic who reside across various governorates, ensuring diverse representation, with primary data collected through standardized and previously validated survey instruments measuring mathematics anxiety, motivation, attitude towards mathematics, and learning strategies. In addition, secondary data were obtained from institutional records, including final mathematics course grades and assessment results, to support the analysis of students' mathematics achievement.

## Participants

The respondents were 227 students enrolled in the Statistics and Data Analysis course at the School of Foundation, Bahrain Polytechnic, during the second semester of AY 2024–2025. Selected from a population of 594 students, they represented learners progressing to Business, Logistics, Web Media, and Visual Design programs, with the course historically recording the lowest passing rate among mathematics courses. The adequacy of the sample size was validated using Soper's (2021) a priori SEM sample size calculator based on an anticipated effect size of 0.30, statistical power of 0.80, five latent variables, twenty-two observed variables, and a 0.05 significance level, which recommended a minimum sample size of 150. The final sample of 227 respondents exceeded this requirement and the recommended minimum for SEM, supporting the reliability and stability of the model estimates.

## Research Instrument

Mathematics anxiety was measured using the Abbreviated Math Anxiety Scale (AMAS), a 9-item Likert-scale instrument consisting of Learning Math Anxiety (5 items) and Math Evaluation Anxiety (4 items), rated from 1 (low anxiety) to 5 (high anxiety) (Hopko et al., 2003). Students' attitudes toward mathematics were assessed using the ATMI-Short, a 20-item instrument measuring self-confidence, value, and enjoyment on a five-point Likert scale (Lim & Chapman, 2013; Tapia & Marsh, 2002). The motivation and learning strategies of students were measured using the Motivated Strategies for Learning Questionnaire (MSLQ), an 81-item instrument covering motivation and learning strategy domains on a seven-point Likert scale (Pintrich et al., 1993). The motivation indicator of ATMI-Short and the test anxiety indicator of MSLQ were excluded since these constructs were measured using separate instruments.

## Data Collection and Procedure

The data collection process was conducted using Google Forms, where all survey instruments, including the AMAS, ATMI, and MSLQ, were consolidated into a single form to ensure simplicity and minimize respondent confusion, while preserving the original item formats and scaling of each validated instrument; clear instructions were provided at the beginning to guide participants. Invitations were sent via institutional email, including a brief description of the study, a unique survey link, and attached informed consent and participant information forms, with a reminder issued midway through the data collection period to improve response rates. Upon completion, all responses were securely downloaded, stored in a password-protected file, and handled with strict confidentiality to prevent unauthorized access.

## Analysis of Data

Mean and summation are used to determine the extent of mathematics proficiency, the level of mathematics anxiety, the level of motivation, and the extent of employing different learning strategies.

Pearson correlation was used to determine the relationships between and among Mathematics achievement and the non-cognitive factors.

Structural Equation Modeling using SPSS AMOS was utilized to identify the best-fit structural equation model of mathematics achievement. Specifically, the analysis evaluated the bivariate relationships among the variables, the direct influence of mathematics anxiety on mathematics achievement, and the mediating role of mathematics anxiety in the relationships between attitude towards mathematics and mathematics achievement. To determine if the model is over-identified, the number of distinct sample moments is calculated using the formula

$$\frac{p(p+1)}{2} \tag{1}$$

and compared it with the number of free parameters. After confirming that the model is over-identified, the data were screened to ensure that key SEM assumptions were satisfied prior to estimation. Multivariate normality, a fundamental requirement for maximum likelihood estimation (MLE), was assessed using Mardia's coefficient of multivariate kurtosis in AMOS, and any violation was addressed through bootstrapping to obtain bias-corrected parameter estimates and robust standard errors, thereby improving the reliability of inference under non-normal conditions (Nevitt & Hancock, 2001). In this study, bootstrapping involved 5,000 replications with replacement while maintaining the original sample size (n = 227), as this number is recommended for stable estimation of confidence intervals and reduced sampling variability (Bollen & Stine, 1990; Efron & Tibshirani, 1993; Nevitt & Hancock, 2001; Preacher & Hayes, 2008). Multicollinearity was assessed using variance inflation factor (VIF), tolerance values, and correlation coefficients among predictors, with values above 0.85 to 0.90 indicating potential issues (Kline, 2016), while outliers were examined using standardized scores with cut-off values of ±3 (Tabachnick & Fidell, 2019). The adequacy of the sample size was confirmed using Soper's (2021) a priori SEM calculator and supported by the minimum recommendation of 200 cases (Boomsma, 1982).

## Ethical Considerations

The study adhered to ethical standards for research involving human participants by obtaining approval from the Research Ethics Committee of Bahrain Polytechnic and the University Research Ethics Committee of Don Mariano Marcos Memorial State University (DMMMSU), along with the necessary institutional permissions. Participation was voluntary, with informed consent obtained from all respondents. Confidentiality and anonymity were ensured by excluding personally identifiable information, anonymizing all data, restricting access to the researcher, and securely storing digital files in accordance with institutional ethical guidelines.

# Results and Discussion

## Mathematics Achievement of Higher Education Students

Table 1 presents the means and passing rate of the respondents enrolled in AP5101 (Statistics and Data Analysis) during Semester 2, AY 2024-2025. Overall, the course achieved a final passing rate of 95.15% (216 out of 227), which is slightly higher than the institutional passing rate of 94.29%.

**Table 1**
*Means and Passing Rate of Respondents Enrolled in AP5101 (Statistics and Data Analysis) Course During Semester 2, AY 2024–2025*

| Assessment | Mean | Pass | % | Fail | % |
|---|---|---|---|---|---|
| Midterm | 74.30 | 181 | 79.74 | 46 | 20.26 |
| Project | 90.19 | 222 | 97.80 | 5 | 2.20 |
| Final | 80.13 | 204 | 89.87 | 23 | 7.93 |
| **Final Grade** | | 216 | 95.15 | 11 | 2.64 |
| **Institutional Passing Rate** | | | **94.29%** | | |

## Level of Non-cognitive Factors of Higher Education Students in Bahrain

### *Level of Mathematics Anxiety*

The overall level of mathematics anxiety is displayed in Table 2. The results show that Learning Math Anxiety was classified as Mild Anxiety with a mean of 2.57, while Math Evaluation Anxiety was rated higher at the Moderate Anxiety level with a mean of 3.12. When combined, the overall mean of 2.82 corresponds to Moderate Anxiety, indicating that students generally experience a manageable but consistent level of worry in mathematics.

**Table 2**
*Overall Level of Mathematics Anxiety*

| Indicator | Mean | Descriptive Level |
|---|---|---|
| 1. Learning Math Anxiety | 2.57 | Mild Anxiety |
| 2. Math Evaluation Anxiety | 3.12 | Moderate Anxiety |
| **Overall** | **2.82** | Moderate Anxiety |

*Note.* Legend: 1.00–1.80 (Low Anxiety), 1.81–2.60 (Mild Anxiety), 2.61–3.40 (Moderate Anxiety), 3.41–4.20 (High Anxiety), 4.21–5.00 (Very High Anxiety)

### *Level of Attitude Towards Mathematics*

Table 3 summarizes the overall level of attitude towards mathematics across the three indicators. The results show that Enjoyment (M = 3.22) and Self-Confidence (M = 3.12) were both rated as Moderate Attitude, while Value obtained the highest mean (M = 3.50) and was classified as a Positive Attitude. When combined, the overall mean score of 3.28 corresponds to a Moderate Attitude.

**Table 3**
*Overall Level of Attitude Towards Mathematics*

| Indicator | Mean | Descriptive Level |
|---|---|---|
| 1. Enjoyment | 3.22 | Moderate Attitude |
| 2. Self-Confidence | 3.12 | Moderate Attitude |
| 3. Value | 3.50 | Positive Attitude |
| **Overall** | **3.28** | Moderate Attitude |

*Note.* Legend: 1.00–1.80 (Very Negative Attitude), 1.81–2.60 (Negative Attitude), 2.61–3.40 (Moderate Attitude), 3.41–4.20 (Positive Attitude), 4.21–5.00 (Very Positive Attitude)

### *Level of Motivation*

Table 4 presents the overall level of motivation of higher education students in Bahrain across five indicators. The results show that Intrinsic Goal Orientation (M = 3.85), Task Value (M = 4.16), and Control of Learning Beliefs (M = 4.18) all fall within the category of Moderate Motivation. On the other hand, Extrinsic Goal Orientation (M = 4.88) and Self-Efficacy for Learning and Performance (M = 4.50) stand out as Slightly High Motivation.

**Table 4**
*Overall Level of Motivation*

| Indicator | Mean | Descriptive Level |
|---|---|---|
| 1. Intrinsic Goal Orientation | 3.85 | Moderate Motivation |
| 2. Extrinsic Goal Orientation | 4.88 | Slightly High Motivation |
| 3. Task Value | 4.16 | Moderate Motivation |
| 4. Control of Learning Beliefs | 4.18 | Moderate Motivation |
| 5. Self-Efficacy for Learning and Performance | 4.50 | Slightly High Motivation |
| **Overall** | **4.34** | Moderate Motivation |

*Note.* Legend: 1.00–1.85 (Very Low Motivation), 1.86–2.71 (Low Motivation), 2.72–3.57 (Slightly Low Motivation), 3.58–4.43 (Moderate Motivation), 4.44–5.29 (Slightly High Motivation), 5.30–6.15 (High Motivation), 6.16–7.00 (Very High Motivation)

### *Level of Learning Strategies*

Table 5 summarizes the overall level of learning strategies across nine indicators. The results show that all components which constitute Rehearsal (M = 4.30), Elaboration (M = 4.26), Organization (M = 4.31), Critical Thinking (M = 4.10), Metacognitive Self-Regulation (M = 4.24), Time and Study Environment (M = 4.34), Effort Regulation (M = 4.42), Peer Learning (M = 4.38), and Help Seeking (M = 4.32) were consistently rated as Moderately Effective Use of Learning Strategies. Among these, Effort Regulation obtained the highest mean, while

Critical Thinking registered the lowest, though all remained within the same descriptive category.

**Table 5**
*Overall Level of Learning Strategies*

| Indicator | Mean | Descriptive Level |
|---|---|---|
| 1. Rehearsal | 4.30 | Moderately Effective Use of Learning Strategies |
| 2. Elaboration | 4.26 | Moderately Effective Use of Learning Strategies |
| 3. Organization | 4.31 | Moderately Effective Use of Learning Strategies |
| 4. Critical Thinking | 4.10 | Moderately Effective Use of Learning Strategies |
| 5. Metacognitive Self-Regulation | 4.24 | Moderately Effective Use of Learning Strategies |
| 6. Time and Study Environment | 4.34 | Moderately Effective Use of Learning Strategies |
| 7. Effort Regulation | 4.42 | Moderately Effective Use of Learning Strategies |
| 8. Peer Learning | 4.38 | Moderately Effective Use of Learning Strategies |
| 9. Help Seeking | 4.32 | Moderately Effective Use of Learning Strategies |
| **Overall** | **4.29** | Moderately Effective Use of Learning Strategies |

*Note.* Legend: 1.00–1.85 (Highly Ineffective Use of Learning Strategies), 1.86–2.71 (Ineffective Use of Learning Strategies), 2.72–3.57 (Slightly Ineffective Use of Learning Strategies), 3.58–4.43 (Moderately Effective Use of Learning Strategies), 4.44–5.29 (Effective Use of Learning Strategies), 5.30–6.15 (Highly Effective Use of Learning Strategies), 6.16–7.00 (Very Highly Effective Use of Learning Strategies)

## Relationship Between and Among Mathematics Achievement and Non-cognitive Factors

Table 6 presents the correlation analysis among mathematics achievement and selected non-cognitive factors, showing several statistically significant relationships at the 0.05 level with varying strengths. Mathematics achievement has a weak positive and statistically significant correlation with motivation (r = 0.374) and learning strategies (r = 0.306), and a moderate positive and statistically significant correlation with attitude towards mathematics (r = 0.498).

In contrast, mathematics anxiety is weakly negative and statistically significantly correlated with mathematics achievement (r = –0.288). Similarly, mathematics anxiety shows weak negative and statistically significant relationships with attitude (r = –0.209) and motivation (r = –0.253), while it has a very weak positive yet statistically significant correlation with learning strategies (r = 0.149). Moreover, attitude towards mathematics had a moderate positive and significant relationship with motivation (r = .433), while its relationship with learning strategies was weak but positive and significant (r = .357). Meanwhile, motivation had a moderate positive and significant relationship with learning strategies (r = .422).

**Table 6**
*Significance of the Relationship Between and Among Mathematics Achievement and Non-cognitive Factors*

| Indicators | Mathematics Achievement | Mathematics Anxiety | Attitude Towards Mathematics | Motivation | Learning Strategies |
|---|---|---|---|---|---|
| Mathematics Achievement | 1 | | | | |
| Mathematics Anxiety | -0.288* | 1 | | | |
| Attitude Towards Mathematics | 0.498* | -0.209* | 1 | | |
| Motivation | 0.374* | -0.253* | 0.433* | 1 | |
| Learning Strategies | 0.306* | 0.149* | 0.357* | 0.422* | 1 |

*Note.* *correlation is significant at 0.05

**Best-Fit Structural Equation Model for Mathematics Achievement as Influenced by Non-cognitive Factors**

The initial model, developed using SPSS Amos, hypothesizes the relationships among five latent constructs influencing mathematics achievement: Attitude Towards Mathematics, Motivation, Learning Strategies, Mathematics Anxiety, and Mathematics Achievement. The model consists of 22 observed indicators and 54 free parameters, with direct paths to Mathematics Achievement and bi-directional correlations among the non-cognitive predictors. With 253 distinct sample moments exceeding the 54 parameters to be estimated, the model is over-identified, allowing the assessment of model fit and the evaluation of the hypothesized relationships.

After confirming that the model was over-identified, the data were screened to ensure that the assumptions for SEM were met. Mardia's coefficient of multivariate kurtosis (128.086; critical ratio = 29.693) indicated a violation of multivariate normality (Byrne, 2016; Kline, 2016), necessitating the use of bootstrapping with 5,000 resamples in AMOS to obtain bias-corrected estimates and robust standard errors.

The bootstrap results demonstrated stable model estimation across 5,000 resamples, with consistent maximum likelihood discrepancy values and low Kullback–Leibler overoptimism. These findings indicate good model stability, low sampling variability, and robust parameter estimation.

To assess multicollinearity, variance inflation factor (VIF) values were all below 10, tolerance values exceeded 0.10, and correlations among the observed variables were below 0.85. These results indicate that multicollinearity was not a concern, allowing reliable estimation of the model parameters (Hair et al., 2020).

The standardized scores of all observed variables were within the acceptable range of ± 3, indicating the absence of problematic outliers. Consequently, all observations were retained for the SEM analysis.

The sample size of 227 was adequate based on Soper's (2021) a priori SEM sample size calculator, which recommended a minimum of 150 participants under the specified model conditions. The sample also exceeded the recommended minimum of 200 cases for reliable SEM analysis (Boomsma, 1982), supporting stable and valid parameter estimates.

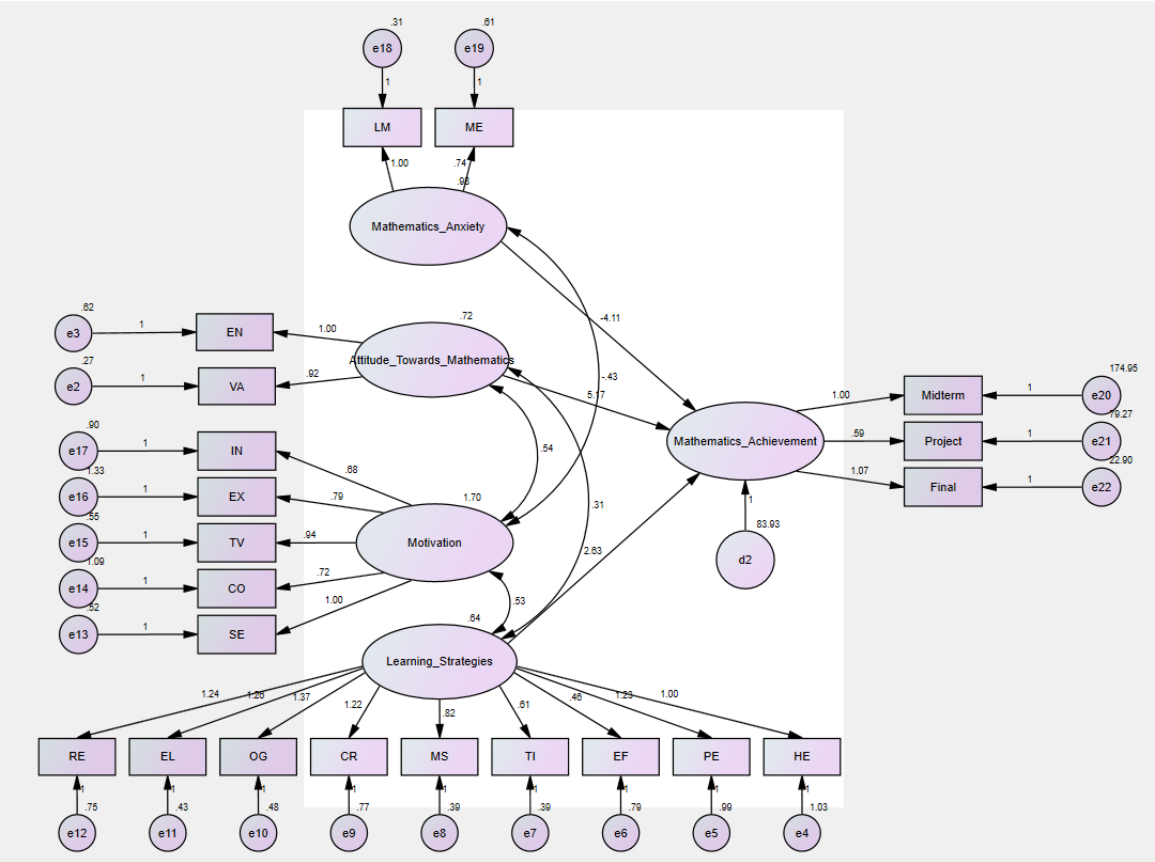
After confirming model identification and assumption testing, all 54 free parameters were estimated simultaneously using the Maximum Likelihood (ML) method with bootstrap corrections to obtain robust parameter estimates under non-normality (Nevitt & Hancock, 2001). All observed indicators loaded significantly on their respective latent constructs except Self-Confidence, which showed a weak and non-significant loading and was therefore removed to improve the measurement model. The remaining indicators demonstrated significant factor loadings, positive variances, and no evidence of Heywood cases, supporting the adequacy of the measurement model.

The structural model showed that Attitude Towards Mathematics, Learning Strategies, and Mathematics Anxiety were significant predictors of Mathematics Achievement, whereas Motivation was not. Consistent with the recommendations of Teo et al. (2013), the non-significant paths were removed to obtain a more parsimonious model. The revised model as shown in Figure 1 consisted of five latent constructs, 21 observed variables, and 51 free parameters, remaining over-identified and suitable for model fit evaluation.

Following the estimation of the hypothesized model, the revised model was evaluated for its adequacy in representing the observed data. All indicators loaded significantly on their respective latent constructs, while the structural model showed that Mathematics Anxiety negatively predicted Mathematics Achievement, whereas Attitude Towards Mathematics and Learning Strategies exerted significant positive effects. These findings indicate that the revised model adequately represents the hypothesized relationships among the constructs.

The initial model demonstrated only a marginally acceptable fit to the data ($\chi^2$ = 551.455, $\chi^2$/df = 2.771, CFI = .853, TLI = .829, RMSEA = .089, SRMR = .0790). After model revision, the fit indices improved ($\chi^2$ = 470.681, $\chi^2$/df = 2.615, CFI = .875, TLI = .854, RMSEA = .085, SRMR = .0661), indicating a more satisfactory representation of the data, although some indices remained below the recommended thresholds.

**Figure 1**
*Revised Mathematics Achievement Model*




Following the recommendations of Byrne (2016) and Kline (2016), model modification was guided by both statistical evidence and theoretical justification. Non-significant covariances between Mathematics Anxiety and Attitude Towards Mathematics (Estimate = –.027, $p$ = .701) and between Mathematics Anxiety and Learning Strategies (Estimate = .072, $p$ = .244) were removed to improve model parsimony. In addition, theoretically justifiable error covariances were added within the Learning Strategies construct (e6–e7, e4–e5, e6–e11, e6–e9, and e6–e8), consistent with recommended SEM practices (Byrne, 2016; Kline, 2016).

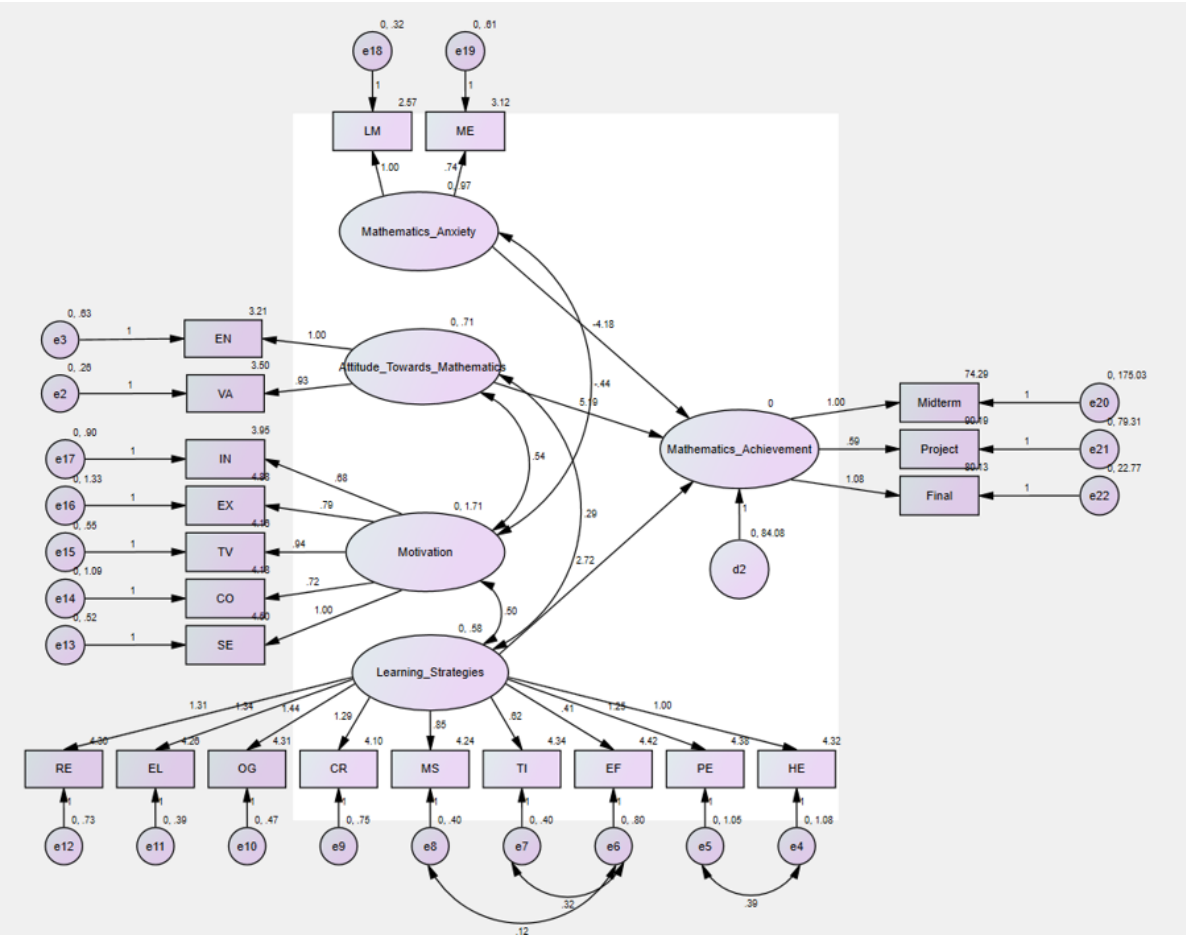
Sequentially adding theoretically justifiable error covariances within the Learning Strategies construct resulted in substantial improvements in model fit. The largest improvement was obtained by adding the covariance between e6 ↔ e7 (M.I. = 77.706), which improved the fit indices to $\chi^2$ = 380.788, $\chi^2$/df = 2.104, CFI = .914, TLI = .900, RMSEA = .070, and SRMR = .0623. Adding e4 ↔ e5 (M.I. = 27.945) further improved the model ($\chi^2$ = 351.608, $\chi^2$/df = 1.953, CFI = .926, TLI = .914, RMSEA = .065, SRMR = .0611), while the addition of e6 ↔ e8 (M.I. = 16.161) produced the final acceptable fit ($\chi^2$ = 337.283, $\chi^2$/df = 1.884, CFI = .932, TLI = .920, RMSEA = .063, SRMR = .0599).

Although additional covariances between e6 ↔ e9 (M.I. = 12.859) and e6 ↔ e11 (M.I. = 10.427) were suggested, they produced only minimal improvements in model fit and were therefore not retained to preserve model parsimony. This decision is consistent with the recommendations of Byrne (2016) and Kline (2016), who emphasized that model modifications should be theoretically justified and should avoid unnecessary model complexity.

The final model as shown in Figure 2 demonstrated a substantially better fit than the revised model. The chi-square decreased from 470.681 ($\chi^2$/df = 2.615) to 337.283 ($\chi^2$/df = 1.884), while SRMR improved from .0661 to .0599 and RMSEA decreased from .085 to .063. Likewise, the incremental fit indices increased from CFI = .875 and TLI = .854 to CFI = .932 and TLI = .920, exceeding the recommended .90 threshold and indicating an acceptable model fit.

The final model consists of five latent constructs and 21 observed variables, with correlated error terms among selected Learning Strategies indicators (e.g., MS–EF, TI–EF, and PE–HE) to account for shared measurement variance.

**Figure 2**
*Final Mathematics Achievement Model*




## Limitations of the Study

This study is limited by its focus on a single institution, Bahrain Polytechnic, which may restrict the generalizability of the findings to other higher education contexts in Bahrain or beyond. Additionally, the reliance on self-reported questionnaire data may introduce response bias, particularly in measuring non-cognitive constructs such as anxiety, motivation, and learning strategies. Although validated instruments were used, their applicability within the specific cultural and institutional context may still pose constraints.

## Conclusion

Based on the findings of the study, it is concluded that students demonstrated strong mathematics achievement, meeting and even exceeding institutional expectations. However, they exhibited only average levels of mathematics anxiety, attitude towards mathematics, motivation, and learning strategy use, indicating that their potential for higher achievement remains underutilized. The significant relationships among these non-cognitive factors suggest that improvements in one area, such as attitude or motivation, can positively influence the others and collectively enhance mathematics performance. Overall, the findings confirm that mathematics achievement is shaped by a combination of mathematics anxiety, attitude towards mathematics, and learning strategies, highlighting the important role of these non-cognitive factors in students' academic success.

In light of the findings, the institution should sustain the effective instructional practices that contributed to student achievement while strengthening support for students' non-cognitive development through evidence-based programs targeting mathematics anxiety, attitudes, motivation, and learning strategies. Integrating effective learning strategies into the curriculum and developing context-specific assessment instruments can further enhance teaching, learning, and institutional decision-making. Future research should involve larger and more diverse samples, examine additional factors influencing mathematics achievement, and ensure instrument validity through pilot testing and confirmatory factor analysis to produce more robust and generalizable findings.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, the author used ChatGPT (OpenAI) solely to assist with language editing, grammar correction, sentence restructuring, and improving the clarity and readability of the text. The AI-assisted tool was not used to generate research ideas, collect or analyze data, interpret results, or formulate the study's conclusions. All aspects of the research, including the study design, methodology, data collection, statistical analyses, interpretation of findings, and final conclusions, were independently conducted and critically reviewed by the author, who assumes full responsibility for the accuracy, originality, and integrity of the work.

## References

Abin, A., Núñez, J. C., Rodríguez, C., Cueli, M., García, T., & Rosário, P. (2020). Predicting Mathematics Achievement in Secondary Education: The role of cognitive, motivational, and emotional variables. *Frontiers in Psychology*, *11*, 876. https://doi.org/10.3389/fpsyg.2020.00876

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-t

Akayuure, F. A., & Akayuure, P. (2024). Examining fresh students' achievement motivation and self-efficacy towards learning high school mathematics. *Contemporary Mathematics and Science Education*, *5*(2), ep24012. https://doi.org/10.30935/conmaths/14691

Alias, Z., Jafar, M. F., & Kasim, M. (2023). The relationship between maths anxiety, attitude towards mathematics and maths problem-solving skills among primary school pupils in kedah. *International Journal of Modern Education, 5*(16), 28–40. https://doi.org/10.35631/ijmoe.516003

Almoslamani, Y. (2022). The impact of learning strategies on the academic achievement of university students in Saudi Arabia. *Learning and Teaching in Higher Education: Gulf Perspectives, 18*(1), 4–18. https://doi.org/10.1108/lthe-08-2020-0025

Ampadu, E., & Anokye-Poku, D. (2022). Influence of personal, motivational and learning environment factors on students' attitudes toward mathematic. *International Journal of Research in Education and Science, 8*(2), 378–392. https://doi.org/10.46328/ijres.2666

Asanre, A. A. (2024). Motivational factors as determinant of mathematics achievement of students in senior secondary schools. *Interdisciplinary Journal of Education Research, 6*, 1–9. https://doi.org/10.38140/ijer-2024.vol6.19

Ashcraft, M. H. (2002). Math Anxiety: Personal, Educational, and Cognitive Consequences. *Current Directions in Psychological Science*, *11*(5), 181–185. https://doi.org/10.1111/1467-8721.00196

Bakar, N. A., Alsmadi, M. S., Ali, Z., Shuaibu, A., & Solahudin, M. H. (2022). Influence of students' motivation on academic achievement among undergraduate students in Malaysia. *Journal of Positive School Psychology, 6*(2), 3443–3450.

Bandura, A. (1985). Model of causality in social learning theory. *Cognition and Psychotherapy*, 81–99. Springer US. https://doi.org/10.1007/978-1-4684-7562-3_3

Baral, D., Dasgupta, N., McLeman, L., Bista, R., Jones, J., & Biswakarma, M. (2025). Exploring the predictive relationship between students' math anxiety and their mathematical performance. *American Journal of STEM Education: Issues and Perspectives*, *3*, 1–15.

Bollen, K. A., & Stine, R. A. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology*, *20*, 115–140. https://doi.org/10.2307/271084

Boomsma, A. (1982). The robustness of LISREL against small sample sizes in factor analysis models. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction* (Part I, pp. 149–173). North-Holland.

Brooks, E. (n.d.). *The Relationship Between Maths Anxiety and Maths Performance*. Centre for Neuroscience in Education. https://www.cne.psychol.cam.ac.uk/relationship-between-maths-anxiety-and-maths-performance

Bryman, A. (2016). *Social Research Methods* (5th ed.). Oxford University Press.

Buckley, S., Reid, K., Goos, M., Lipp, O. V., & Thomson, S. (2016). Understanding and addressing mathematics anxiety using perspectives from education, psychology and neuroscience. *Australian Journal of Education, 60*(2), 157–170. https://doi.org/10.1177/0004944116653000

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming, third edition*. Routledge. https://doi.org/10.4324/9781315757421

Caputol, A. R., & Chan, L. (2024). Attitude towards mathematics as a mediator in the relationship between motivated strategies for learning and anxiety. *Journal Healthcare Treatment Development, 45*, 1–13. https://doi.org/10.55529/jhtd.45.1.13

Caviola, S., Mammarella, I. C., & Kovas, Y. (2019). Math anxiety in children with and without mathematical difficulties. *Mathematics Anxiety*, 141–155. Routledge. https://doi.org/10.4324/9780429199981-8

Caviola, S., Toffalini, E., Giofrè, D., Ruiz, J. M., Szűcs, D., & Mammarella, I. C. (2022). Math Performance and Academic Anxiety Forms, from Sociodemographic to Cognitive Aspects: A Meta-analysis on 906,311 Participants. *Educational Psychology Review, 34*(1), 363–399. https://doi.org/10.1007/s10648-021-09618-5

Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). Sage Publications.

Davidovitch, N., & Dorot, R. (2023). The effect of motivation for learning among high school students and undergraduate students—A comparative study. *International Education Studies, 16*(2), 117–127. https://doi.org/10.5539/ies.v16n2p117

Dowker, A., & Sheridan, H. (2022). Relationships between mathematics performance and attitude to mathematics: Influences of gender, test anxiety, and working memory. *Frontiers in Psychology, 13*, 814992. https://doi.org/10.3389/fpsyg.2022.814992

Earp, J. (2024). PISA 2022 insights: Student questions, mathematics anxiety and homework. *Teacher Magazine*. https://doi.org/10.3679.80

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64*(3), 830. https://doi.org/10.2307/1131221

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Springer US. https://doi.org/10.1007/978-1-4899-4541-9

Fennema, E. H., & Sherman, J. A. (1978). Sex-related differences in mathematics achievement and related factors: A further study. *Journal for Research in Mathematics Education, 9*(3), 189–203. https://doi.org/10.2307/748997

Fernández-Cézar, R., Solano-Pinto, N., & Garrido, D. (2021). Can Mathematics Achievement Be Predicted? The Role of Cognitive–Behavioral–Emotional Variables. *Mathematics*, *9*(14), 1591. https://doi.org/10.3390/math9141591

García-Pérez, D., Fraile, J., & Panadero, E. (2020). Learning strategies and self-regulation in context: how higher education students approach different courses, assessments, and challenges. *European Journal of Psychology of Education*, *36*(2), 533–550. https://doi.org/10.1007/s10212-020-00488-z

Gursel, G., Övez, F., & Öztürk, G. (2023). High school students' level of mathematics anxiety and use of self-regulatory learning strategies. *Acta Didactica Napocensia, 16*, 170–181. https://doi.org/10.24193/adn.16.1.12

Ha, C., Roehrig, A. D., & Zhang, Q. (2023). Self-regulated learning strategies and academic achievement in South Korean 6th-graders: A two-level hierarchical linear modeling analysis. *PLoS ONE*, *18*(4), e0284385. https://doi.org/10.1371/journal.pone.0284385

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2020). *Multivariate data analysis* (8th ed.). Cengage Learning.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21*, 33–46.

Hopko, D., Mahadevan, R., Bare, R., & Hunt, M. (2003). The Abbreviated Math Anxiety Scale (AMAS). *Assessment, 10*(2), 178–182. https://www.oecd.org/en/publications/pisa-2022-results-volume-i_53f23881-en.html

Hwang, S., & Son, T. (2021). Students' Attitude toward Mathematics and its Relationship with Mathematics Achievement. *Journal of Education and E-Learning Research*, *8*(3), 272–280. https://doi.org/10.20448/journal.509.2021.83.272.280

Jehadus, E., Tamur, M., Chen, J., & Perbowo, K. S. (2022). The influence of tutoring and learning motivation on mathematics achievement of junior high school students. *Journal of Honai Math*, *5*(1), 75–82. https://doi.org/10.30862/jhm.v5i1.206

Khan, B. S. B., & Salman, R. (2020). Influence of mathematics in our daily lives. *Arts & Humanities Open Access Journal*, *4*(2), 50–52.

Kiwanuka, H. N., Van Damme, J., Van den Noortgate, W., & Reynolds, C. (2020). Temporal relationship between attitude toward mathematics and mathematics achievement. *International Journal of Mathematical Education in Science and Technology*, *53*(6), 1546–1570. https://doi.org/10.1080/0020739X.2020.1832268

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford Publications.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.

Kumar, N. A. (2020). INVESTIGATION INTO MOTIVATION: SIGNIFICANCE, TYPES AND THEORIES. *EPRA International Journal of Economics Business and Management Studies*, 44–49. https://doi.org/10.36713/epra5993

Lens, W., & Vansteenkiste, M. (2020). Motivation: About the "why" and "what for" of human behavior. In *Psychology Press eBooks* (pp. 249–270). https://doi.org/10.4324/9781003076384-12

Libradilla, A. A., Morales, D. R. B., Piamonte-Bacatan, R. J., & Misoles, P. A. (2023). Examining the association between attitude towards mathematics and student achievement. *International Journal of Research Publication and Reviews, 4*(7), 1040–1046.

Liebendörfer, M., Göller, R., Gildehaus, L., Kortemeyer, J., Biehler, R., Hochmuth, R., Ostsieker, L., Rode, J., & Schaper, N. (2022). The role of learning strategies for performance in mathematics courses for engineers. *International Journal of Mathematical Education in Science and Technology, 53*(5), 1133–1152. https://doi.org/10.1080/0020739x.2021.2023772

Lim, S., & Chapman, E. (2013). Development of a short form of the attitudes toward mathematics inventory. *Educational Studies in Mathematics, 82*, 145–164. https://doi.org/10.1007/s10649-012-9414-x

Marticion, R. D. (2021). Mathematical anxiety and learning motivation strategies among senior high school students in the STEM program. *Southeast Asian Mathematics Education Journal, 11*(1), 13–24.

Mendaje, J. R. C. (2018). *Performance of sophomore secondary students exposed in the Kumon mathematics program*. *World Journal of Research and Review, 6*(3), 68–74.

Mohr-Schroeder, M. J., Jackson, C., Cavalcanti, M., Jong, C., Craig Schroeder, D., & Speler, L. G. (2017). Parents' attitudes toward mathematics and the influence on their students' attitudes toward mathematics: A quantitative study. *School Science and Mathematics, 117(*5), 214–222. https://doi.org/10.1111/ssm.12225

Molnár, G., & Kocsis, Á. (2023). Cognitive and non-cognitive predictors of academic success in higher education: a large-scale longitudinal study. *Studies in Higher Education*, *49*(9), 1610–1624. https://doi.org/10.1080/03075079.2023.2271513

Moss, N. E., & Moss-Racusin, L. (2021). Academic achievement. In *Best practices in child and adolescent behavioral health care* (pp. 85–91). https://doi.org/10.1007/978-3-030-73515-9_13

Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling*, *8*(3), 353–377. https://doi.org/10.1207/S15328007SEM0803_2

OECD. (2024). *PISA 2022 Results (Volume V): Learning Strategies and Attitudes for Life.* PISA, OECD Publishing. https://doi.org/10.1787/c2e44201-en

Okudan, Ü., & Yeşilyurt, E. (2023). Learning Needs, General Teaching Principles, and Learning Strategies as Predictors of Mathematics Academic Achievement. *The Asia-Pacific Education Researcher*. https://doi.org/10.1007/s40299-023-00715-3

Oluyomi, J., Salami, T., & Yomi, A. (2024). The influence of mathematics anxiety, motivation, and attitudes on academic performance among senior secondary school students in Katsina State, Nigeria. *Mediterranean Publications and Research International: International Journal of Assessment & Evaluation in Education*, *5*(8), 153–168.

Ovella, K. (2023). *Assessing math motivation, educational expectancy, and grades in corequisite coursework to predict outcomes in college algebra* [Master's thesis, University of New Orleans]. University of New Orleans Theses and Dissertations. https://scholarworks.uno.edu/td/3109

Pintrich, P. R., Smith, D. A. F., Garcia, T., & Mckeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (mslq). *Educational and Psychological Measurement, 53*(3), 801–813. https://doi.org/10.1177/0013164493053003024

Pizon, M., & Ytoc, S. (2021). A Path Model to Infer Mathematics Performance: The Interrelated Impact of Motivation, Attitude, Learning Style and Teaching Strategies Variables. *East Asian Journal of Multidisciplinary Research, 1*(3), 315–330. https://doi.org/10.48550/arXiv.2105.05850

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879–891. https://doi.org/10.3758/BRM.40.3.879

Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: Psychometric data. *Journal of Counseling Psychology, 19*(6), 551–554. https://doi.org/10.1037/h0033456

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68.

Salkind, N. J. (2017). *Exploring research* (9th ed.). Pearson Education.

Sangaram, S., & Maat, S. M. (2020). Systematic Literature Review on Secondary School Students Mathematics Anxiety. *The International Journal of Academic Research in Business and Social Sciences, 10*, 1347–1363.

Saqr, M., Matcha, W., Ahmad Uzir, N., Jovanovic, J., Gašević, D., & López-Pernas, S. (2023). Transferring effective learning strategies across learning contexts matters: A study in problem-based learning. *Australasian Journal of Educational Technology*, *39*(3), 35–57. https://doi.org/10.14742/ajet.8303

Schukajlow, S., Rakoczy, K., & Pekrun, R. (2023). Emotions and motivation in mathematics education: Where we are today and where we need to go. *ZDM*, *55*(2), 249–267. https://doi.org/10.1007/s11858-022-01463-2

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology, 60*, 101832. https://doi.org/10.1016/j.cedpsych.2019.101832

Seagull, B. (2024, August 23). *Maths education is failing UK students*. @FinancialTimes; Financial Times. https://www.ft.com/content/c2c0f389-324f-412d-9e16-ba6d3031dbb1?

Sengül, S., & Dereli, M. (2013). The effect of learning integers using cartoons on 7th grade students' attitude to mathematics. *Educational Sciences: Theory and Practice, 13*(4), 2526–2534.

Sharma, P. (2021). Importance and application of mathematics in everyday life. *International Journal for Research in Applied Science & Engineering Technology*, *9*(XI), 868–879. https://doi.org/10.22214/ijraset.2021.38869

Shear, B. R. (2020). Mathematics achievement. In F. Maggino (Ed.), *Encyclopedia of quality of life and well-being research* (pp. 1–3). Springer. https://doi.org/10.1007/978-3-319-69909-7_1746-2

Shum, A., & Fryer, L. K. (2023). Grade goal effects on the interplay between motivation and performance in undergraduate gateway mathematics courses. *Contemporary Educational Psychology, 75*, 102228. https://doi.org/10.1016/j.cedpsych.2023.102228

Siller, H.-S., & Ahmad, S. (2024). Analyzing the impact of collaborative learning approach on grade six students' mathematics achievement and attitude towards mathematics. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(2), em2395. https://doi.org/10.29333/ejmste/14153

Soper, D. S. (2021). *A-priori sample size calculator for structural equation models*. https://www.danielsoper.com/statcalc

Strunk, K. K., & Mwavita, M. (2024). *Design and Analysis in Quantitative Educational Research: Univariate Designs in SPSS* (2nd ed.). Routledge.

Supriadi, N., Jamaluddin Z, W., & Suherman, S. (2024). The role of learning anxiety and mathematical reasoning as predictor of promoting learning motivation: The mediating role of mathematical problem solving. *Thinking Skills and Creativity, 52*, 101497. https://doi.org/10.1016/j.tsc.2024.101497

Süren, N., & Kandemir, M. A. (2020). The effects of mathematics anxiety and motivation on students' mathematics achievement. *International Journal of Education in Mathematics, Science and Technology, 8*(3), 190. https://doi.org/10.46328/ijemst.v8i3.926

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Tapia, M., & Marsh, G. E., II. (2002). Confirmatory factor analysis of the Attitudes Toward Mathematics Inventory. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Chattanooga, TN. (ERIC Document Reproduction Service No. ED 471 301).

Telaumbanua, A. R., Harefa, A. O., & Zega, Y. (2023). Analysis of Students' Concept Understanding Ability in View of Mathematical Anxiety in the Implementation of Learning in Class VII of SMP Negeri 1 Mazo. *QISTINA Jurnal Multidisiplin Indonesia*, *2*(1), 177–188. https://doi.org/10.57235/qistina.v2i1.564

Teo, T., Tsai, L. T., & Yang, C.-C. (2013). Applying structural equation modeling (SEM) in educational research. *Application of Structural Equation Modeling in Educational Research and Practice*, 3–21. SensePublishers. https://doi.org/10.1007/978-94-6209-332-4_1

Tuncer, M., & Yilmaz, Ö. (2020). Relations attitude towards mathematics lesson, anxiety and academic success. *Journal of Research in Mathematics Education, 9*(2), 173. https://doi.org/10.17583/redimat.2020.4061

Villa, E. A., & Sebastian, M. A. (2021). Achievement motivation, locus of control and study habits as predictors of mathematics achievement of new college students. *International Electronic Journal of Mathematics Education, 16*(3), em0661. https://doi.org/10.29333/iejme/11297

von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U., & Yin, L. (2024). *TIMSS 2023 International Results in Mathematics and Science*. Boston College, TIMSS & PIRLS International Study Center. https://doi.org/10.6017/lse.tpisc.timss.rs6460

Wanamaker, C. (2024, March 13). *The importance of mathematics education in global society*. HubPages. https://discover.hubpages.com/education/the-importance-of-mathematics-education-in-an-increasingly-complex-modern-global-society

Wen, R., & Dubé, A. K. (2022). A systematic review of secondary students' attitudes towards mathematics and its relations with mathematics achievement. *Journal of Numerical Cognition, 8*(2), 295–325. https://doi.org/10.5964/jnc.7937

Wild, S., & Neef, C. (2023). Analyzing the associations between motivation and academic performance via the mediator variables of specific mathematic cognitive learning strategies in different subject domains of higher education. *International Journal of STEM Education, 10*(1). https://doi.org/10.1186/s40594-023-00423-w

Yip, M. C. W. (2019). *Learning strategies*. Oxford Bibliographies in Education. https://doi.org/10.1093/obo/9780199756810-0211

Zhang, J., Zhao, N., & Kong, Q. P. (2019). The relationship between math anxiety and math performance: A meta-analytic investigation. *Frontiers in Psychology, 10*. https://doi.org/10.3389/fpsyg.2019.01613

# Migration Intentions Among Mathematics Educators in Higher Education Institutions (HEIs) in Batangas, Philippines

Lea Catapang, Bahrain Polytechnic, Bahrain

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This study examined the migration intentions of mathematics educators in higher education institutions (HEIs) in Batangas, Philippines, during Academic Year 2025–2026. Using a descriptive-correlational design, data were gathered from 150 mathematics educators through a researcher-developed questionnaire. Descriptive statistics, Spearman's rank correlation, Chi-square tests with Cramér's V, and Partial Least Squares Structural Equation Modeling (PLS-SEM) were employed to examine the relationships among institutional, individual, and external factors; job satisfaction; selected demographic characteristics; and migration intention. Results revealed that individual and external factors emerged as the strongest predictors of migration intention, whereas institutional factors, job satisfaction, and demographic variables exerted comparatively weaker influence. Notably, while educators reported strong consideration of working abroad, their corresponding preparatory actions remained only moderate, pointing to a discernible gap between aspiration and action. The structural model accounted for 29.3% of the variance in migration intention, indicating meaningful explanatory power despite the multifaceted nature of migration decision-making. These findings suggest that migration intention among mathematics educators is shaped less by dissatisfaction with current institutional conditions and more by external opportunities and personal aspirations for growth. Accordingly, HEIs and policymakers are encouraged to strengthen career advancement pathways, provide greater global exposure, and design targeted retention strategies that respond to educators' professional ambitions, thereby mitigating the risk of workforce attrition in mathematics education.

*Keywords:* migration intention, mathematics educators, higher education institutions, PLS-SEM

iafor
The International Academic Forum
www.iafor.org

## Introduction

The international migration of educators has become an increasingly significant challenge for higher education systems worldwide. Rather than representing a simple response to job dissatisfaction, migration is now recognized as a complex career decision shaped by economic opportunities, institutional conditions, professional aspirations, and broader structural forces within a globalized academic labor market (Haddad et al., 2021). The migration of mathematics and other STEM educators is particularly concerning because these disciplines are fundamental to scientific innovation, technological advancement, and national development. Persistent teacher shortages in developed countries have intensified international recruitment from developing nations, accelerating global competition for qualified educators. UNESCO (2024) estimates that an additional 44 million teachers will be needed worldwide by 2030, with teacher attrition accounting for more than half of this shortage, thereby increasing international demand for experienced educators.

The Philippines is deeply embedded in this global migration system. Teacher migration has evolved from occasional labor mobility into a sustained pattern of overseas employment driven by both economic opportunities and professional advancement. National statistics indicate that approximately 1,500 teachers migrated annually for overseas employment between 2013 and 2017 (Philippine Statistics Authority [PSA], 2023), reflecting a continuing trend rather than isolated cases. The CALABARZON region, one of the country's leading sources of overseas Filipino workers, accounted for 16.35% of all land-based deployments during the third quarter of 2024 (PSA, 2023). Within this region, Batangas Province serves as an important educational hub, hosting 63 higher education institutions (Commission on Higher Education [CHED], 2024). However, local institutions have increasingly experienced the loss of mathematics educators to overseas employment, particularly in the post-pandemic period, creating challenges for faculty retention, institutional continuity, and long-term educational quality.

Previous research has consistently identified institutional and economic conditions as major determinants of educator migration. Low salaries, heavy teaching workloads, limited career advancement, and inadequate organizational support act as important push factors, whereas higher compensation, stronger research environments, professional recognition, and better career opportunities abroad serve as powerful pull factors (Alicamen & Becamon, 2022; Narvaza, 2023; Yaw, 2022). Beyond institutional influences, individual aspirations, family considerations, and long-term career goals also shape migration decisions. Job satisfaction has similarly been associated with migration intentions, with higher satisfaction generally reducing educators' willingness to seek employment abroad (Ersoy, 2024; Malinao & Agustin, 2023). Demographic characteristics, particularly age, family status, and employment status, further influence migration decisions by affecting individuals' willingness and capacity to relocate internationally (Dennison, 2022; Dibeh et al., 2020; Iştekli, 2025).

This study is primarily grounded in the Push–Pull–Mooring (PPM) framework (Moon, 1995), which explains migration as the result of interacting forces that encourage departure, attract individuals to alternative destinations, and either facilitate or constrain movement. The framework is complemented by Lee's Migration Theory (Lee, 1966), Human Capital Theory (Becker, 1964; Sjaastad, 1962), Herzberg's Two-Factor Theory (Herzberg et al., 1959), the Theory of Planned Behavior (Ajzen, 1991), Network Theory (Massey et al., 1993), the New Economics of Labor Migration (Stark & Bloom, 1985), Job Embeddedness Theory (Mitchell et al., 2001), and Life Course Theory (Elder, 1994). Together, these perspectives provide a

comprehensive explanation of how institutional conditions, personal motivations, social relationships, and economic opportunities interact to shape educators' migration intentions. Accordingly, this study conceptualizes migration intention as being influenced by four latent constructs: institutional factors, individual factors, external factors, and job satisfaction.

Despite the growing literature on teacher migration, several important gaps remain. First, relatively few studies focus specifically on mathematics educators despite their critical role in strengthening STEM education and national competitiveness. Second, Philippine studies have largely employed descriptive or qualitative approaches, limiting their ability to examine the complex relationships among multiple determinants of migration intentions. Advanced multivariate techniques such as Partial Least Squares Structural Equation Modeling (PLS-SEM) enable simultaneous examination of these interrelationships (Hair et al., 2017; Henseler et al., 2009). Third, empirical evidence from provincial higher education contexts such as Batangas Province remains limited despite increasing migration pressures experienced by local institutions.

To address these gaps, this study employed PLS-SEM to examine how institutional factors, individual factors, external factors, and job satisfaction influence migration intentions among mathematics educators in higher education institutions in Batangas Province, Philippines. By integrating multiple theoretical perspectives within a single structural model, the study contributes to the growing literature on educator mobility while providing evidence to inform faculty retention strategies and higher education workforce policy.

Specifically, this study sought to answer the following research questions:

1. What is the demographic profile of mathematics educators in higher education institutions in Batangas Province in terms of age, family status, and employment status?
2. What is the extent of migration intentions among these educators?
3. What is the level of perceived influence of institutional factors, individual factors, external factors, and job satisfaction on migration intentions?
4. Is there a significant relationship between demographic characteristics, institutional factors, individual factors, external factors, job satisfaction, and migration intentions?
5. What structural model best explains migration intentions among mathematics educators in higher education institutions in Batangas Province?

## Methods

This study employed a quantitative descriptive–correlational research design to examine the determinants of migration intentions among mathematics educators. The descriptive component profiled institutional, individual, and external factors, together with job satisfaction, whereas the correlational component examined the relationships between these predictors and migration intention without manipulating any variables (Creswell & Creswell, 2018). The quantitative approach enabled objective measurement of observed and latent constructs through standardized instruments and statistical analyses. The design was particularly appropriate for Partial Least Squares Structural Equation Modeling (PLS-SEM), which permits the simultaneous analysis of multiple interrelated predictors of migration intention.

The study population comprised mathematics educators employed in Commission on Higher Education (CHED)-recognized higher education institutions in Batangas Province,

Philippines. Eligible participants were full-time, contractual, and part-time faculty members teaching mathematics or mathematics-related courses who had at least one year of higher education teaching experience. Faculty members serving exclusively in administrative positions, teaching outside mathematics, on extended leave, or employed outside Batangas Province were excluded.

Sample size was estimated using power analysis recommendations for PLS-SEM (Hair et al., 2017) assuming 80% statistical power, $\alpha = .05$, a medium effect size ($f^2 = .15$), and four predictors. The minimum required sample was 85 respondents. To account for clustering of educators within institutions, the estimate was adjusted using a design effect of 1.35, resulting in a target sample of 115 respondents.

The questionnaire measured institutional factors, individual factors, external influences, job satisfaction, and migration intention using multi-item latent constructs derived from the International Teacher Mobility Survey (Caravatti et al., 2014) and the Job Satisfaction Survey (Spector, 1985). All substantive items were measured on a five-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). Items were randomized to minimize order effects and response bias (Stefkovics & Kmetty, 2022).

Content validity was established through two rounds of expert review following Lynn (1986). The final instrument achieved an I-CVI of 1.00 and an S-CVI/Ave of 1.00, exceeding recommended thresholds (Polit & Beck, 2006). A pilot study involving 10 mathematics educators demonstrated satisfactory internal consistency (Cronbach's $\alpha = .861–.896$; McDonald's $\omega = .866–.901$). Reliability and validity were reassessed using the final study sample during PLS-SEM analysis.

Data collection was carried out via an online structured questionnaire administered through Google Forms. Invitations were distributed using institutional and personal email addresses following a multi-wave strategy to maximize response rates (Dillman et al., 2014). The researcher also personally visited selected HEIs to facilitate survey administration and enhance participation through direct engagement. Prior to completing the questionnaire, all respondents were provided with a written informed consent form outlining the study's purpose, scope, and procedures, as well as their rights to decline participation or withdraw at any time without penalty. Ethical clearance was obtained from the Institutional Research Ethics Committee before data collection commenced, and all procedures complied with the Data Privacy Act of 2012.

Data were analyzed using R. Descriptive statistics summarized respondents' demographic characteristics and the distributions of all study variables. Spearman's rank-order correlation assessed bivariate associations among the latent constructs because the questionnaire responses were ordinal. Associations between demographic variables and migration intention were examined using chi-square tests of independence and Cramér's V. The hypothesized relationships among latent variables were subsequently evaluated using Partial Least Squares Structural Equation Modeling (PLS-SEM), including assessment of the measurement model and the structural model.

# Results

## Demographic Profile of Mathematics Educators

Table 1 presents the demographic profile of the 150 mathematics educators who participated in the study. Respondents were predominantly mid-career professionals, with a mean age of 39.59 years (SD = 8.80) and the largest proportion falling within the 36–40 age bracket (25.33%). Early-career entrants (25–30 years old) comprised only 12.67% of the sample, while those under 24 or over 56 represented minimal proportions, reflecting a generally mature and professionally established workforce. Most respondents were married with children (42.0%) or single (32.0%), and the majority held full-time permanent positions (64.0%).

**Table 1**
*Demographic Profile of Mathematics Educators*

| Profile | Frequency | Percentage |
|---|---|---|
| ***Age*** | | |
| ≤ 24 years old | 2 | 1.33 |
| 25-30 years old | 19 | 12.67 |
| 31-35 years old | 29 | 19.33 |
| 36-40 years old | 38 | 25.33 |
| 41-45 years old | 31 | 20.67 |
| 46-50 years old | 14 | 9.33 |
| 51-55 years old | 10 | 6.67 |
| 56-60 years old | 1 | 0.67 |
| ≥ 61 years old | 6 | 4.00 |
| ***Family Status*** | | |
| Single | 48 | 32.00 |
| Married, no children | 13 | 8.67 |
| Married, with children | 63 | 42.00 |
| Single parent | 15 | 10.00 |
| Divorced/Separated | 6 | 4.00 |
| Widowed | 5 | 3.33 |
| ***Employment Status*** | | |
| Full-time Permanent | 96 | 64.00 |
| Full-time Temporary | 14 | 9.33 |
| Full-time Contractual | 25 | 16.67 |
| Part-time | 15 | 10.00 |

*Source*: Original Research Results

## Extent of Migration Intentions

Table 2 presents the extent of migration intentions across three progressive indicators reflecting cognitive consideration, exploratory behavior, and preparatory action. Educators strongly agreed that they were seriously considering working abroad, indicating high migration intention on this indicator. Exploratory and preparatory behaviors, however, both registered only moderate intention, yielding an overall moderate migration intention rating. This pattern reveals a notable divergence between educators' cognitive readiness to migrate and their actual behavioral activation—an aspiration–action gap.

**Table 2**
*Extent of Migration Intentions Among Mathematics Educators*

| Indicators | MFR | DR | Interpretation |
|---|---|---|---|
| I am seriously considering working in another country soon. (MI1) | 4 | A | High MI |
| I have taken steps to explore professional opportunities abroad. (MI2) | 3 | NA nor D | Moderate MI |
| I am preparing the requirements needed for overseas employment. (MI3) | 3 | NA nor D | Moderate MI |
| **Overall** | **3** | **NA nor D** | **Moderate MI** |

*Source*: Original Research Results. *Note:* MFR = Most Frequent Response. Scale: 5 = Strongly Agree, 4 = Agree, 3 = Neither Agree nor Disagree, 2 = Disagree, 1 = Strongly Disagree.

## Level of Perceived Influence of Factors on Migration Intentions

Table 3 summarizes the perceived influence of institutional, individual, external, and job satisfaction factors on migration intention. Individual and external factors showed uniformly high perceived influence across all respective sub-indicators.

**Table 3**
*Level of Perceived Influence of Factors on Migration Intentions*

| Factor/Indicator | MFR | DR | Interpretation |
|---|---|---|---|
| ***Institutional*** | | | |
| Teaching Workload (INSF1) | 4 | A | High perceived influence on MI |
| Administrative Support (INSF2) | 3 | NA nor D | Moderate perceived influence on MI |
| Work Environment (INSF3) | 4 | A | High perceived influence on MI |
| Organizational Culture (INSF4) | 3 | NA nor D | Moderate perceived influence on MI |
| **Overall** | **4** | **A** | **High perceived influence on MI** |
| ***Individual*** | | | |
| Career Advancement Goals (INDF1) | 4 | A | High perceived influence on MI |
| Family Welfare (INDF2) | 4 | A | High perceived influence on MI |
| Long-Term Goals (INDF3) | 4 | A | High perceived influence on MI |
| Cultural Adaptability (INDF4) | 4 | A | High perceived influence on MI |
| **Overall** | **4** | **A** | **High perceived influence on MI** |
| ***External*** | | | |
| Salary Differentials (EXTF1) | 4 | A | High perceived influence on MI |
| Professional Recognition (EXTF2) | 4 | A | High perceived influence on MI |
| Research Opportunities (EXTF3) | 4 | A | High perceived influence on MI |
| Social Network (EXTF4) | 4 | A | High perceived influence on MI |
| **Overall** | **4** | **A** | **High perceived influence on MI** |
| ***Job Satisfaction*** | | | |
| Teaching Satisfaction (JS1) | 4 | A | High perceived influence on MI |
| Emotional Fulfilment (JS2) | 3 | NA nor D | Moderate perceived influence on MI |
| Stress Level (JS3) | 4 | A | High perceived influence on MI |
| Institutional Pride (JS4) | 3 | NA nor D | Moderate perceived influence on MI |
| **Overall** | **4** | **A** | **High perceived influence on MI** |

*Source*: Original Research Results

Institutional factors and job satisfaction each presented a mixed pattern: teaching workload and work environment, and teaching satisfaction and stress level, were rated as highly influential, whereas administrative support, organizational culture, emotional fulfilment, and institutional pride were rated as only moderately influential. Despite this internal variation, all four factor categories registered a high overall rating, indicating a generally strong perceived influence on migration intention.

## Relationships Between Constructs and Migration Intentions

Table 4 presents the relationships between the study constructs and migration intention. Spearman's rank correlation revealed statistically significant positive associations between migration intention and external factors ($\rho = .429$, $p < .001$), individual factors ($\rho = .382$, $p < .001$), and institutional factors ($\rho = .214$, $p < .01$). Job satisfaction, by contrast, showed a significant negative association ($\rho = -.191$, $p < .05$), suggesting that lower job satisfaction corresponds with greater migration intention. Chi-square tests with Cramér's V indicated that none of the demographic variables—age, family status, and employment status—were significantly associated with migration intention, and these variables were accordingly excluded from the structural model.

**Table 4**
*Relationships Between Constructs and Migration Intentions*

| Construct | ρ | p-value |
|---|---|---|
| Institutional Factors | 0.214** | 0.009 |
| Individual Factors | 0.382*** | 0.000 |
| External Factors | 0.429*** | 0.000 |
| Job Satisfaction | -0.191* | 0.019 |
| Age | 0.249 | 0.246 |
| Family Status | 0.162 | 0.731 |
| Employment Status | 0.155 | 0.545 |

*Source*: Original Research Results. *Note*: *** ($p < .001$), ** ($p < .01$), * ($p < .05$)

## PLS-SEM Results

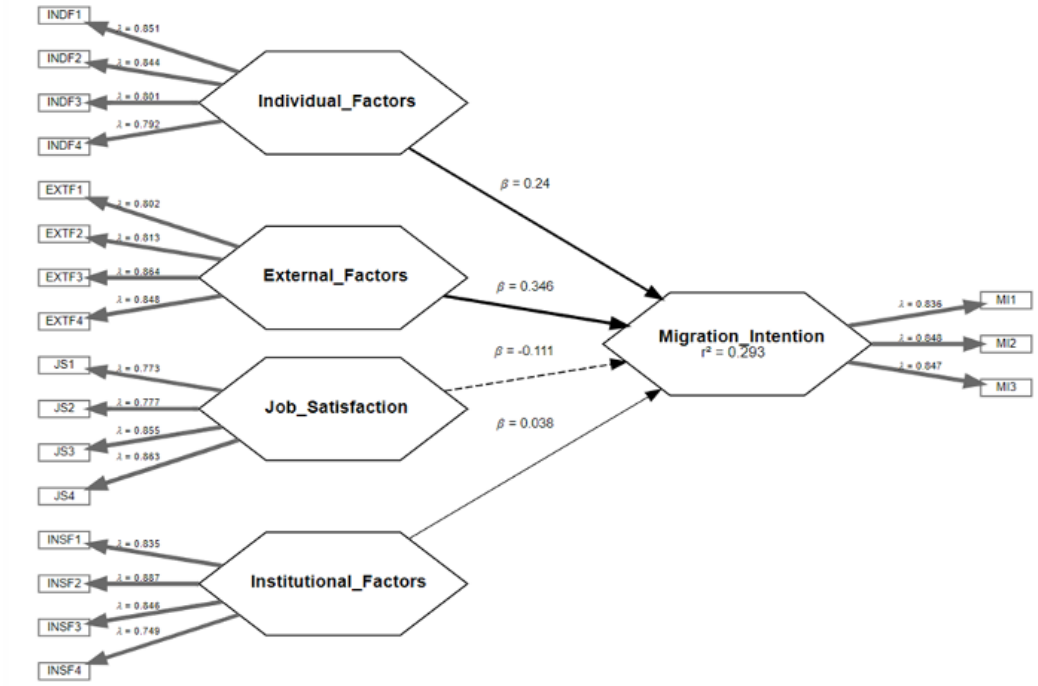
Given the absence of statistically significant associations between the demographic variables and migration intention, age, family status, and employment status were excluded from the structural model.

The measurement model (Figure 1) demonstrated strong psychometric properties. All indicator loadings ranged from 0.779 to 0.867, exceeding the recommended 0.70 threshold, while variance inflation factor (VIF) values ranged from 1.847 to 2.520, well below the conservative ceiling of 5.0, confirming the absence of problematic multicollinearity. Internal consistency reliability was supported by Cronbach's alpha values ranging from 0.798 to 0.857, composite reliability (ρc) ranging from 0.881 to 0.902, and rhoA ranging from 0.799 to 0.877, all within the acceptable 0.70–0.95 range. Convergent validity was established, with Average Variance Extracted (AVE) values exceeding 0.50 for all constructs (0.674–0.712). Discriminant validity was confirmed through the Fornell–Larcker criterion, with the square roots of AVE (0.821–0.844) exceeding corresponding inter-construct correlations, and

through Heterotrait-Monotrait (HTMT) ratios ranging from 0.236 to 0.558, all below the conservative 0.85 threshold.

The structural model (Figure 1) explained 29.3% of the variance in migration intention ($R^2$ = 0.293; Adjusted $R^2$ = 0.274). Given the multifaceted and context-dependent nature of migration decision-making, this level of explained variance represents substantial explanatory power (Cohen, 1988).

**Figure 1**
*Structural Model of Migration Intentions Among Mathematics Educators in HEIs*




*Source*: Original Research Results

## Discussion

### Demographic Characteristics and Their Contextual Role

The demographic profile of mathematics educators in Batangas higher education institutions provides important context for understanding migration intentions. The predominance of mid-career educators suggests a workforce that is actively evaluating long-term professional growth, financial stability, and career advancement, consistent with studies indicating that educators at this career stage are more likely to reassess employment opportunities in pursuit of better prospects (Tschopp et al., 2016). Furthermore, the substantial proportion of married respondents with children highlights that migration decisions are shaped not only by individual aspirations but also by household considerations. This finding supports the New Economics of Labor Migration, which conceptualizes migration as a family-based strategy to improve economic security and diversify household income rather than as an exclusively individual decision (Stark & Bloom, 1985).

Although most respondents held permanent appointments, employment security alone did not appear sufficient to discourage migration intentions. Previous research suggests that stable employment may reduce uncertainty but does not necessarily offset the influence of limited

compensation, heavy workloads, or constrained career advancement opportunities (Castellacci & Viñas-Bardolet, 2021). Thus, permanent employment should not be interpreted as an assurance of workforce retention when more attractive professional opportunities exist elsewhere.

Despite these demographic patterns, age, family status, and employment status were not significantly associated with migration intentions. This finding indicates that demographic characteristics primarily provide the context within which migration decisions are made rather than directly determining educators' intentions to migrate. Consistent with the Push–Pull–Mooring framework (Moon, 1995), these characteristics function as mooring factors that may influence the timing or feasibility of migration without serving as its principal drivers. Consequently, migration intentions appear to be distributed across the mathematics educator workforce rather than concentrated within particular demographic groups. This suggests that faculty retention initiatives should address the broader institutional and professional conditions influencing migration instead of focusing solely on specific demographic segments.

**Extent of Migration Intentions: A Pre-migration State**

The disparity between high cognitive intention and only moderate behavioral preparation suggests that many mathematics educators are in a pre-migration state, where migration is being seriously considered but has not yet progressed to concrete action. This finding is consistent with Lee's (1966) distinction between migration aspirations and actual migration behavior and supports the Push–Pull–Mooring (PPM) framework (Moon, 1995), which proposes that mooring factors—such as family responsibilities, institutional ties, and practical constraints—may delay or inhibit the translation of migration intentions into action. Similar intention–behavior gaps have been reported among Filipino teachers (Tantay et al., 2024) and healthcare professionals (Ohene-Botwe et al., 2025), suggesting that this transitional stage is characteristic of skilled professionals in developing countries.

The moderate overall level of migration intention should therefore not be interpreted as evidence of workforce stability. Rather, it reflects a transitional state in which educators remain open to migration but have not yet initiated concrete plans. According to the Theory of Planned Behavior (Ajzen, 1991), behavioral intentions may rapidly develop into action when favorable opportunities arise and perceived barriers are reduced. Consequently, external events such as attractive overseas employment offers, the migration of professional peers, changes in institutional policies, or worsening working conditions may accelerate the progression from contemplation to active migration (Soto Nishimura & Czaika, 2024).

These findings highlight the importance of timely institutional intervention. Retention strategies are likely to be most effective while educators remain in this pre-migration stage, before migration intentions evolve into concrete job search behaviors and relocation plans. Proactive measures that enhance career development opportunities, improve working conditions, and strengthen organizational commitment may therefore reduce the likelihood that migration aspirations translate into actual workforce losses.

**Perceived Influence of Migration Determinants**

The consistently high perceived influence of institutional, individual, external, and job satisfaction factors indicates that mathematics educators view migration decisions as the

product of multiple interacting influences rather than a single dominant cause. This finding supports the multidimensional nature of migration proposed by the Push–Pull–Mooring (PPM) framework (Moon, 1995), in which migration intentions emerge from the combined effects of workplace conditions, personal aspirations, external opportunities, and contextual constraints.

Among the institutional factors, teaching workload and work environment were perceived as more influential than administrative support and organizational culture. This pattern is consistent with Herzberg's Two-Factor Theory (Herzberg et al., 1959), suggesting that immediate working conditions exert a stronger influence on migration considerations than broader organizational characteristics. It also reflects evidence that excessive job demands and limited workplace resources contribute to professional strain and increase intentions to seek employment elsewhere (Nair et al., 2022; Narvaza, 2023).

The uniformly high ratings across all individual and external factors indicate that migration is driven not only by dissatisfaction with current employment but also by proactive career aspirations. Career advancement, family welfare, long-term goals, salary differentials, professional recognition, research opportunities, and overseas professional networks were all perceived as important influences, reinforcing previous findings that skilled migration is motivated by both personal ambitions and the attractiveness of international opportunities (Becker, 1964).

A more nuanced pattern emerged for job satisfaction. While respondents reported high influence for teaching satisfaction and work-related stress, emotional fulfillment and institutional pride were rated only moderately. This suggests that educators remain strongly committed to the profession of teaching but feel less attached to their employing institutions. Such partial organizational attachment is consistent with Job Embeddedness Theory (Mitchell et al., 2001), which proposes that weaker organizational connections reduce the costs of leaving an institution. Consequently, educators who continue to value teaching yet perceive greater professional opportunities abroad may be more willing to transfer their expertise to better-resourced educational systems without experiencing substantial professional or personal disruption.

Overall, these findings indicate that migration intentions among mathematics educators arise from the interaction of institutional conditions, individual aspirations, global opportunities, and organizational attachment. Faculty retention efforts should therefore extend beyond improving workplace conditions alone and incorporate strategies that strengthen career development, professional recognition, research support, and institutional commitment. Addressing only one dimension of the migration process is unlikely to reduce migration intentions when multiple influences operate simultaneously.

**Relationships Between Migration Determinants and Migration Intention**

The significant positive correlations between institutional factors ($\rho = .214$), individual factors ($\rho = .382$), external factors ($\rho = .429$), and migration intentions indicate that mathematics educators' migration intentions are associated with multiple interrelated influences rather than a single determinant. Among these, external factors demonstrated the strongest association, suggesting that opportunities such as higher salaries, professional recognition, research opportunities, and established overseas networks are closely linked to educators' intentions to migrate. This finding is consistent with migration theories

emphasizing the influence of external pull factors (Harris & Todaro, 1970; Lee, 1966) and with previous studies reporting that international career opportunities remain a major driver of skilled migration (Ohene-Botwe et al., 2025; Zulfiu Alili et al., 2023).

Job satisfaction exhibited a significant negative correlation with migration intentions, indicating that educators who reported higher levels of job satisfaction were generally less inclined to consider overseas employment. This finding supports previous evidence that job satisfaction serves as a protective factor against turnover and migration intentions (Ersoy, 2024). However, the relatively weak correlation suggests that improving job satisfaction alone may be insufficient to offset the influence of stronger external opportunities and individual career aspirations. This finding reinforces the view that migration decisions result from the combined influence of workplace experiences and opportunities beyond the employing institution.

In contrast, age, family status, and employment status were not significantly associated with migration intentions. This finding suggests that demographic characteristics primarily provide the context within which migration decisions are made rather than directly influencing educators' intentions to migrate. Consistent with the Push–Pull–Mooring framework (Moon, 1995), demographic characteristics may shape the feasibility or timing of migration without acting as independent determinants of migration intention.

Taken together, these findings reinforce the multidimensional nature of educator migration. Migration intentions appear to be associated with the interaction of institutional conditions, individual aspirations, external opportunities, and job satisfaction rather than any single factor in isolation. While the correlation analysis identifies the strength and direction of these relationships, the structural model provides a more comprehensive assessment of the independent contribution of each construct when considered simultaneously.

**Structural Model: External and Individual Factors as Primary Drivers**

The PLS-SEM analysis provides a more comprehensive understanding of migration intentions by examining the unique contribution of each predictor while controlling for the effects of the others. The structural model explained 29.3% of the variance in migration intention ($R^2 = .293$), indicating substantial explanatory power for a behavioral outcome influenced by multiple personal, organizational, and contextual factors. The unexplained variance suggests that additional influences, such as individual psychological characteristics, recruitment opportunities, professional experiences, or broader economic conditions, may also contribute to migration decisions and warrant further investigation.

Among the four latent constructs, only external factors ($\beta = .346$, $p < .001$; $f^2 = .142$) and individual factors ($\beta = .240$, $p < .001$; $f^2 = .068$) exerted significant direct effects on migration intention. The stronger influence of external factors highlights the importance of salary differentials, professional recognition, research opportunities, and overseas networks in shaping educators' intentions to migrate. This finding supports migration theories that emphasize the role of external pull forces in skilled migration (Harris & Todaro, 1970; Lee, 1966) and is consistent with previous studies identifying international career opportunities as a major determinant of educator mobility (Ohene-Botwe et al., 2025). The significant effect of individual factors further suggests that migration is not solely a response to unfavorable institutional conditions but also reflects proactive career planning, family aspirations, and

long-term professional goals, consistent with Human Capital Theory (Becker, 1964) and the Theory of Planned Behavior (Ajzen, 1991).

In contrast, institutional factors and job satisfaction did not demonstrate statistically significant direct effects after accounting for the influence of the other predictors. Although both constructs were significantly associated with migration intentions in the bivariate analysis, their effects were no longer significant in the structural model, suggesting that their influence may be shared with or overshadowed by stronger external and individual determinants. These findings should not be interpreted as evidence that institutional conditions and job satisfaction are unimportant. Rather, they indicate that, within the present model, they do not independently explain migration intentions beyond the effects of external opportunities and individual motivations. Their inclusion remains theoretically justified because migration decisions are widely recognized as multidimensional processes involving organizational, personal, and environmental influences (Hair et al., 2021; Lee, 1966; Mitchell et al., 2001).

The contrast between the correlation and structural model results provides an important theoretical contribution. While institutional factors and job satisfaction are associated with migration intentions, external opportunities and individual aspirations emerge as the most influential predictors when all factors are considered simultaneously. This finding suggests that improving workplace conditions alone may not be sufficient to reduce migration intentions if substantial disparities in career opportunities, professional recognition, and financial rewards continue to exist between domestic and international academic environments. Consequently, effective faculty retention strategies should extend beyond institutional improvements to include stronger career development pathways, enhanced research support, competitive compensation, and broader opportunities for professional growth.

Future research should examine whether institutional factors and job satisfaction influence migration intentions indirectly through mediating or moderating mechanisms. Longitudinal studies would also provide valuable insights into how educators' migration intentions evolve over time and whether changes in institutional conditions alter the relative influence of external opportunities and individual aspirations.

## Conclusion

The findings demonstrate that migration intentions among mathematics educators are not determined by a single factor but emerge from the interplay of personal aspirations, institutional experiences, and external opportunities. Although institutional conditions and job satisfaction remain important considerations, the stronger influence of external opportunities and individual motivations suggests that retaining mathematics educators requires more than improving the workplace environment alone. As global demand for qualified educators continues to grow, higher education institutions must adopt proactive and comprehensive retention strategies that support career advancement, research productivity, professional recognition, and competitive employment conditions. Ultimately, strengthening the capacity of institutions to retain mathematics educators is essential not only for sustaining a high-quality teaching workforce but also for ensuring the long-term development of mathematics education and higher education in the Philippines.

## Acknowledgements

This article is based on the author's doctoral dissertation. The author wishes to thank her dissertation adviser and oral examination panel for their guidance and valuable feedback during the conduct of this study, as well as the colleges and universities in Batangas Province, their administrations, and mathematics faculty members who supported and participated in the research.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Claude, an AI-assisted language model developed by Anthropic, was used solely for proofreading and refining the language of the manuscript, limited to correcting grammatical errors and improving clarity. All ideas, analyses, and findings are the author's own original work.

## References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Alicamen, D. B. L., & Becamon, A. M. K. M. (2022). Overseas teaching: The voice of Filipino early childhood education teachers in Singapore. *Pertanika Journal of Social Science and Humanities, 30*(4), 1473–1494. https://doi.org/10.47836/pjssh.30.4.02

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. National Bureau of Economic Research & Columbia University Press.

Caravatti, M., McLeod Lederer, S., Lupico, A., & Van Meter, N. (2014). *Getting teacher migration & mobility right*. Education International.

Castellacci, F., & Viñas-Bardolet, C. (2021). Internet use and job satisfaction. *Computers in Human Behavior, 118*, Article 106718. https://doi.org/10.1016/j.chb.2021.106718

Commission on Higher Education. (2024). *List of higher education institutions*. https://ched.gov.ph/list-of-higher-education-institutions-2/

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dennison, J. (2022). Re-thinking the drivers of regular and irregular migration: Evidence from the MENA region. *Comparative Migration Studies, 10*, Article 21. https://doi.org/10.1186/s40878-022-00296-y

Dibeh, G., Fakih, A., & Marrouch, W. (2020). Decision to emigrate amongst youth in developing countries: Evidence from the Middle East. *International Migration, 58*(3), 1–19. https://doi.org/10.1111/imig.12639

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method* (4th ed.). Wiley.

Elder, G. H., Jr. (1994). Time, human agency, and social change: Perspectives on the life course. *Social Psychology Quarterly, 57*(1), 4–15. https://doi.org/10.2307/2786971

Ersoy, A. (2024). The relationship between job satisfaction and turnover intention: A meta-analysis. *Journal of Human and Work, 11*(2), 115–124. https://doi.org/10.18394/iid.1464432

Haddad, B., Milliken, D. B., Stewart, J., & Velez, J. J. (2021). Then what? Quantifying SBAE teacher career-decisions post-migration. *Journal of Agricultural Education, 62*(1), 131–143. https://doi.org/10.5032/jae.2021.01131

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). SAGE Publications. https://doi.org/10.4135/9781526470383

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). SAGE Publications. https://doi.org/10.1007/978-3-030-80519-7

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. In R. R. Sinkovics & P. N. Ghauri (Eds.), *Advances in International Marketing* (Vol. 20, pp. 277–319). Emerald Group Publishing Limited. https://doi.org/10.1108/S1474-7979(2009)0000020014

Herzberg, F., Mausner, B., & Snyderman, B. B. (1959). *The motivation to work* (2nd ed.). John Wiley & Sons.

İstekli, M. (2025). Global migration dynamics: The driving force of poverty and connections and the impact of demographic factors. *İktisat Politikası Araştırmaları Dergisi, 12*(1), 68–84. https://doi.org/10.26650/JEPR1548794

Lee, E. S. (1966). A theory of migration. *Demography, 3*(1), 47–57. https://doi.org/10.2307/2060063

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382–385.

Malinao, C. W., & Agustin, J. (2023). Retention, satisfaction, and motivation enablers of a higher education institution in the Philippines amidst teacher migration. *Puissant, 4*, 716–735.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., & Taylor, J. E. (1993). Theories of international migration: A review and appraisal. *Population and Development Review, 19*(3), 431–466. https://doi.org/10.2307/2938462

Mitchell, T. R., Holtom, B. C., Lee, T. W., Sablynski, C. J., & Erez, M. (2001). Job embeddedness: Using job embeddedness to predict voluntary turnover. *Academy of Management Journal, 44*(6), 1102–1121. https://doi.org/10.2307/3069391

Moon, B. (1995). Paradigms in migration research: Exploring “moorings” as a schema. *Progress in Human Geography, 19*(4), 504–524. https://doi.org/10.1177/030913259501900404

Nair, S., Lim, Y. M., & Cheik, A. N. (2022). Internal push and external pull factors affecting lecturers' turnover intention in Malaysian private universities. *Behavioral Sciences, 12*(7), Article 236. https://doi.org/10.3390/bs12070236

Narvaza, M. (2023). Job security and migration intentions of higher education faculty in the Philippines. *Asian Education Studies, 8*(4), 30–49.

Ohene-Botwe, B., Antwi, W. K., Amedu, C., Akudjedu, T. N., Mudadi, L. S-., Chinene, B., Adjeley Quaye, S. N., Dambele, M. Y., Mkoloma, S. S., Barare, C., Kumsa, M. J., Sichone, J., Saizi, R., Mdletshe, S., Dlama, J. Z., Malamateniou, C., & Ahmad, R. (2025). Challenges and strategies for retaining Africa's radiography workforce in the continent amidst a competitive global market: Insights from 17 countries. *Radiography, 31*(4), Article 103001. https://doi.org/10.1016/j.radi.2025.103001

Philippine Statistics Authority. (2023, January 16). *Migration and overseas workers: 2020 census of population and housing (2020 CPH).*

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*(5), 489–497.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy, 70*(5), 80–93. https://doi.org/10.1086/258726

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). SAGE Publications.

Soto Nishimura, A., & Czaika, M. (2024). Exploring migration determinants: A meta-analysis of migration drivers and estimates. *Journal of International Migration and Integration, 25*, 621–643. https://doi.org/10.1007/s12134-023-01091-z

Spector, P. E. (1985). Measurement of human service staff satisfaction: Development of the Job Satisfaction Survey. *American Journal of Community Psychology, 13*(6), 693–713. https://doi.org/10.1007/BF00929796

Stark, O., & Bloom, D. E. (1985). The new economics of labor migration. *The American Economic Review, 75*(2), 173–178.

Stefkovics, Á., & Kmetty, Z. (2022). A comparison of question order effects on item-by-item and grid formats: Visual layout matters. *Measurement Instruments for the Social Sciences, 4*, Article 8. https://doi.org/10.1186/s42409-022-00036-z

Tantay, M. A., et al. (2024). Teachers' motivations, challenges, and insights on working abroad: A qualitative exploration. *Harbin Gongcheng Daxue Xuebao (Journal of Harbin Engineering University), 45*(6), 67–72.

Tschopp, C., Grote, G., & Koppel, N. (2016). Disentangling effects of age and career preferences on the relationship between job satisfaction and turnover intention and behavior: An examination in three samples. *Work, Aging and Retirement, 2*, 73–85. https://doi.org/10.1093/workar/wav022

UNESCO & International Task Force on Teachers for Education 2030. (2024). *Global report on teachers: Addressing teacher shortages and transforming the profession*. UNESCO.

Yaw, J. (2022). *Professional emigration and the impact of globalization on developing countries: "Brain drain" in the Philippines' education sector* (Master's thesis). Graduate School of International Studies, Seoul National University.

Yaw, R. (2022). Organizational culture and teacher migration: A case from Southeast Asia. *International Journal of Educational Mobility, 5*(3), 49–64.

Zulfiu Alili, M., King, R., & Memeti, M. (2023). The migration intentions of university staff in North Macedonia: A potential brain drain? *Eastern European Economics, 62*(5), 633–649. https://doi.org/10.1080/00128775.2023.2225494

# Technology-Mediated English Grammar Learning by Chinese Learners: A Case Study

Chi Wai Lee, Hong Kong Metropolitan University, Hong Kong

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Traditional paper-based in-class writing tasks are common when we teach students English grammar in universities. Nowadays, technology including artificial intelligence (AI) can assist students' learning. Using an experiential approach to learning, this study firstly attempts to reveal the extent to which technology-mediated grammar learning across course-topics can be realised. Secondly, it aims to report students' evaluation of this approach of technology-mediated grammar learning. Methodologically, this study embraces a learning model of "experiencing, reflecting, thinking and acting." It integrates technology tools and student-centred activities of a course "English Grammar." At the term-end, 30 year-2 students posted their responses to open-ended questions onto the University's online forum. The technology tools include (a) an in-house online platform, (b) Open AI, c) websites and d) YouTube. The hands-on practices cover topics: (i) thesis statements, (ii) effective writing, (iii) functions of writing, and (iv) parts of speech. Students use their smartphones to (a) post paragraphs onto a university's online forum when elaborating on a topic, (b) use OpenAI to generate myths of effective writing and discuss their flaws, (c) search for fundraiser websites and evaluate functions of writing, and (d) browse YouTube and explore parts of speech in song lyrics. The above has revealed technology-mediated grammar learning across topics can be successfully realised through experiential learning. Students' qualitative comments regarding this integrated approach are: 1) paper-pencil writing tasks are perceived as "old-fashioned," 2) authenticity and interaction through technology meets their expectations, 3) the development of AI causes them to question the significance of grammar learning.

*Keywords:* technology-mediated grammar learning, experiential learning, L2 English

iafor
The International Academic Forum
www.iafor.org

## Introduction

When it comes to teaching and learning English grammar in higher education, core topics of the course are clearly stated in a course syllabus. Traditionally, those topics include parts of speech, topic sentences and supporting ideas, functions and effectiveness of writing, etc. In today's world, technology develops rapidly, particularly with the prevalent use of Artificial Intelligence (AI), and this definitely has an impact on how our students should learn English grammar through technology. In this study, the first aim is to examine the extent to which technology-mediated grammar learning can be implemented across course topics. That is, course topics can be well-designed to be taught by integrating different technology tools into teaching and learning English grammar; this purposefully embraces both action research approach and an experiential learning theory. The second aim is to present students' qualitative feedback on this integrated approach combining technology-mediated and language learning.

Over the past decades, action research has evolved to incorporate digital technologies, participatory approaches, and applications in multilingual and AI-enhanced learning environments. Action research is widely recognised as a practitioner-driven, cyclical methodology aimed at improving professional practice through systematic inquiry. Originating with Kurt Lewin (1946), who conceptualised research as iterative cycles of planning, action, and evaluation. Anne Burns (1999, 2010, 2019) positions action research in education discipline as a tool for teacher development and classroom innovation. Contemporary studies (Bradbury, 2022; Paulus & Wise, 2019; Stringer, & Aragón, 2021) highlight its value for empowering practitioners, generating context-specific knowledge, and addressing complex educational and social challenges. To comply with action research approach, experiential learning is also key in education. According to Kolb (1984), "experiential learning is the process whereby knowledge is created through the transformation of experience" (p. 41). Beard and Wilson (2013) also state that

> experiential learning often has formalized processes where people take time out to reflect on what has happened during an activity.... In effect, formal and experiential learning processes have varying elements of each other and learning might be considered more of a continuum of authenticity. (p. 39)

In this study, the two approaches of action research (see Burns 2010, 2019) and experiential learning (see Beard & Wilson 2013; Boyatzis & Mainemelis 2001; Kolb & Kolb 2009; Kolb, Malinen 2000) are adopted. Specifically, the following two aims of this study are presented below:

1. To examine the extent to which technology-mediated grammar learning can be implemented across course topics.
2. To present students' qualitative feedback on this integrated approach combining technology-mediated and language learning.

## Literature Review

### Experimental Learning

To begin, Hutton (1989) states that "experiential learning is learning that is rooted in our doing and our experience. It is learning which illuminates that experience and provides direction for the making of judgments as a guide to choice and action" (p. 51). Beard and

Wilson (2013) explain that "The insight gained through the conscious and unconscious internationalization of our own or observed experiences which build upon our past experiences or knowledge" (p. 16; see Norma, 2010; among many others). In experiential learning theory proposed by Kolb and Kolb (2009), there are major propositions that the foundational experiential learning scholars shared. Those major propositions are briefly quoted (2009) as follows.

Learning:

1. is best conceived as a process, not in terms of outcomes. To improve learning in higher education, the primary focus should be on engaging students in a process that best enhances their learning.
2. is best facilitated by a process that draws out the students' beliefs and ideas about a topic so that they can be examined, tested and integrated with new, more refined ideas.
3. requires the resolution of conflicts between dialectically opposed modes of adaptation to the world…. In the process of learning one is called upon to move back and forth between opposing modes of reflection and action and feeling and thinking.
4. is a holistic process of adaptation. It is not just the result of cognition but involves the integrated functioning of the total person – thinking, feeling, perceiving and behaving.
5. results from synergetic transactions between the person and the environment. (pp. 43-44)

Kolb and Kolb (2009) further elaborate the experiential learning model as (see McLeod 2017)

> an idealized learning cycle or spiral where the learner 'touches all the bases' – experiencing, reflecting, thinking and acting – in a recursive process that is responsive to the learning situation and what is being learned. Immediate or concrete experiences are the basis for observations and reflections. (p. 44)

To conclude, this section has highlighted those key notions of experiential learning derived by Kolb and Kolb (2009). The notions include: Concrete Experience, Abstract Conceptualization, Reflective Observation and Active Experimentation. When taking those four core notions of an experiential learning framework into account, the method has embedded active experimentation and reflective observation by specific task design and interview questions, embracing the notions of abstract conceptualization and concrete experience.

## Methodology

This section is to explain the two methods with specific task design, targeting at achieving the two aims of this study correspondingly.

### Method 1: Action-Based and Experiential Learning Approach

This study firstly aims to examine the extent to which technology-mediated grammar learning can be implemented across course topics. As such, this study has purposefully employed four designated technology tools for students in class, and this is basically to have embraced the framework of "experiential learning" – learning from doing (Kolb & Kolb, 2009). Specifically, given four different topics of the course titled "English grammar," 30 students in class were carefully guided to employ the four specific technology tools when learning the

following course topics. Students in class were expected to use their smartphone or tablet when achieving the following tasks.

***Task 1***

- Technology tools: in-house online learning platform named "OLE"
- Focused course topic: A topic sentence and supporting ideas

- Activity: In class, students are expected to use their smartphones or iPads to: (a) post "paragraphs in English" onto an in-house online learning platform "OLE" after selecting some random topics such as:
    - What supernatural power would you like to acquire? Why?
    - What would you do if you were at the end-of-life stage and why?
    - What colour would you like to represent you? Why?

***Task 2***

- Technology tools: University's supported Azure Open AI
- Focused course topic: Myths of English writing

- Activity: Students are expected to use the University's Azure OpenAI to and then discuss:

  To generate 10 myths about English writing and critique their shortcomings, for instance,
    - You must never spilt an infinitive.
    - You must never end a sentence with a preposition.

***Task 3***

- Technology tools: online websites
- Focused course topic: functions of writing – informative, persuasive, etc.

- Activity: Students are expected to use their smartphones or iPads to: c) examine crowdfunding platforms (e.g. kickstarter.com) to assess clarity and persuasive writing functions in texts targeting potential backers
    - To what extent would those selected written descriptions of a product or service strongly appeal to you as a potential backers (as an investor)?
    - Comment on those words or sentences that particularly draw your attention. Why?

  Fundraising websites (e.g. kickstarter.com) are introduced to students, and students are asked to browse any projects with descriptions in English that appeal to them.

***Task 4***

- Technology tools: YouTube
- Focused course topic: Parts of speech

- Activity: Students are expected to use their smartphones or iPads to: (d) watch YouTube film songs, analysing lyrics with respective to parts of speech

  Song types include:
  - Aladdin – Prince Ali Lyrics
  - The Princess and the Frog

  Watch a video clip (Aladdin – Prince Ali Lyrics) and find out different parts of speech/phrasal structures

**Method 2: Qualitative Student Feedback Through On-line Discussion Platform**

To collect data for addressing the second aim of this study, at the end of the term, the cohort of 30 students was invited to voluntarily submit their reflections on the integrated technology-mediated grammar-learning approach via the University's online discussion platform. It should be noted that the sample size is not large, but it is worth investigating and collecting narrative responses from students who experienced the use of all four technology tools namely in-house online learning platform, university-supported Open AI tools, websites and YouTube.

The goal of this method is to collect the users' perceived values when integrating technology into learning English grammar; it is not about evaluating the effectiveness of this teaching and learning method. Those four discussion questions include:

- Do you favour paper-based grammar exercises over technology-mediated tools such as apps or websites in classroom learning? Please explain.
- Do you agree that artificial intelligence (AI) supports learners in producing grammatically accurate English?
- In AI era, how important do you consider English Grammar Course for students?
- Any other comments?

## Results and Discussion

### Implementation of Technology-Mediated L2 English Grammar Learning

***Task 1 – Our In-House Online Learning Platform Named "OLE"***

Student Sample Work: Topic – What supernatural power would you like to acquire? Why?

If I can have a supernatural power, it will be teleportation. I like to travel around the world but the flight tickets are always expensive. If I possess the ability to teleport to wherever I want, I could save a bunch of money. Moreover, this ability can be very useful in daily life too. I won't have to worry about being late since I can just teleport. No need to go out early for traffic, and save more time for myself.

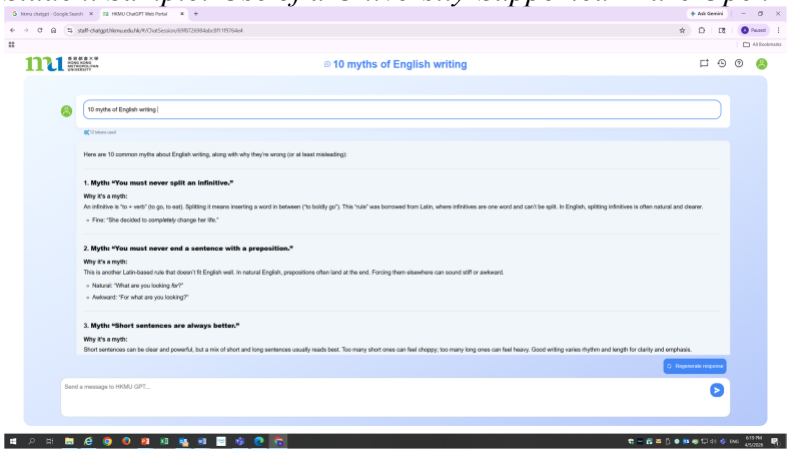
***Task 2 – University's Supported Azure Open AI***

The following figure is a student sample work when using a university-supported Azure Open AI. Students in class refer to the feedback shown by OpenAI and students discuss whether they fully agree to, particularly with respect to traditional grammar and modern usage approaches to writing.

Feedback from Open AI: Rule 1: You must never spilt an infinitive. This rule was borrowed from Latin, where infinitives are one word and can't be split.

**Figure 1**

*Student Sample: Use of a University-Supported Azure OpenAI Service*




### *Task 3 – Online and Authentic Websites*

In class, students are introduced the informative and persuasive functions of writing, and they are trained to learn how to pitch in English so as to attract readers' attention. Traditionally, students are given a writing task and then asked to write on paper for completing a selling task. When integrating technology into teaching and learning this course topic, students were guided to explore authentic fundraising websites and to investigate how appealing a real product or service was. Specifically, one of the authentic fundraising websites among many others is "kickstarter.com" and it is a platform from which the public can read those written descriptions of business products or services. Students in class were expected to find out if the public can be highly convinced to be a "backer" who is willing to back up this business idea. For example, the following is one extract from many business ideas (source: kickstarter.com):

***Why This Product?***

- LOCKLERY comes as a six-piece set of two knives, two spoons and two forks, all made of high quality stainless steel and FDA grade materials, designed in an elegant shape
- LOCKLERY is also eco-friendly, using sustainable polymers like Polypropylene (PP) and Thermoplastic Rubber (TPR), making it a responsible choice for the environmentally conscious design lover

Students were required to give critical comments on those English descriptions such as the uses of what adjectives, concrete or mass nouns which could lead to a small or big number of backers, with respect to the writing style or other visual communication strategies, etc.

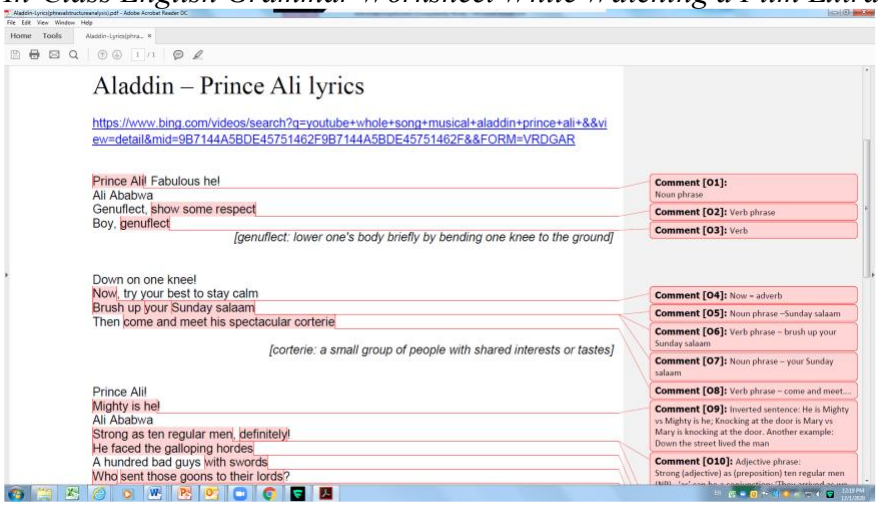
### *Task 4 – YouTube: Film Extracts*

Students in class watched a film extract with song and its lyrics, and they were expected to complete an English grammar worksheet – identifying parts of speech and phrasal structures in English (see Figure 2) while enjoying a film extract. The choice of film songs vary and

they can be musical genre such as Aladdin – Prince Ali lyrics or other genres which students would love to watch.

**Figure 2**

*In-Class English Grammar Worksheet While Watching a Film Extract*




*Note.* https://www.bing.com/videos/riverview/relatedvideo?q=youtube%20aladdin%20songs&mid=F417180850603F37FECBF417180850603F37FECB&ajaxhist=0

To conclude, the above section has clearly showcased the implementation of those L2 English grammar learning tasks via the uses of four specific technology tools namely in-house online learning platform, university-supported Open AI tool, websites and YouTube.

## Students' Qualitative Feedback on This Technology-Mediated L2 English Learning

To recap, there are four discussion questions (see below) for our students, and they aim to collect qualitative feedback on this technology-mediated L2 English learning through the University's online learning discussion platform. Out of 30 respondents, the respondent rate is 67% (20).

1. Do you favour paper-based grammar exercises over technology-mediated tools such as apps or websites in classroom learning? Please explain.
2. Do you agree that artificial intelligence (AI) supports learners in producing grammatically accurate English?
3. In AI era, how important do you consider English Grammar Course for students?
4. Any other comments?

### *Students' Representative Direct Quotes*

> It's boring to only use paper-pen to write in class. It's old fashioned!
>
> The use of technology appeals to most of us because I can type and browse through some apps and websites. Class time seems to move fast.
>
> It's more authentic to use websites or apps or AI when working out our in-class exercises.

> The use of technology when learning grammar should be promoted because we prefer not to sit and write in class.
>
> AI is very useful when writing anything, and their English is of course better than our standard.
>
> I like using AI to correct my writing and I found that they have clear explanations on what is wrong and what is correct.
>
> AI is great! It's quick to produce good English. My concern is that I cannot write properly when there is no AI available.
>
> English grammar learning is just one subject and I just want to get better scores.
>
> AI develops superfast, we can simply ask AI to write and I don't think English grammar class is important.

Concluding the above section, the results of qualitative feedback have revealed that: a) paper-based writing tasks are perceived as "old fashioned," b) authenticity and interaction through technology tools meet students' expectations, and c) the development of AI causes them to question about the significance of grammar learning. Specifically, students responded that they perceive that implementing technology-mediated L2 English grammar learning is what they expect, and traditional paper-pen writing in class is counted as boring or old fashioned. Students reveal that they would use their smartphones or tablets to browse authentic materials from the real-world, rather than from textbooks. All in all, most students accepted this kind of hands-on experiential learning approach to learning. It should be also noted that some students' qualitative feedback cast some doubts on significance of their learning English grammar in university due to rapid AI development.

## Conclusion

In summary, this qualitative study has clearly pointed to the results that students' experimental learning of L2 English grammar via numerous technology tools can be implemented across different course topics. Conventional paper-based writing approach in class should not be counted as an only way of teaching English grammar. Clearly, technology-mediated English grammar learning is well-accepted by our target university students and the technology tools include in-house university's learning platform, Open AI, websites and YouTube, etc. Throughout this study, students' hands-on experiential learning with the effective use of smartphones or tablets in class can be demonstrated. Students' qualitative comments have also revealed that: a) paper-based writing tasks are perceived as "old fashioned," b) authenticity and interaction through technology while learning are expected, and c) casting doubts on significance of students' grammar learning due to AI development is found. Admittedly, the sample size of this study is small and the findings have limited generalizability.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that no AI or AI-assisted technologies have been used to generate, refine or correct the content in the manuscript.

## References

Beard, C., & Wilson, J. (2013). *Experiential learning*. Kogan Page.

Bradbury, H. (2022). *The SAGE handbook of action research*. SAGE.

Burns, A. (1999). *Collaborative action research for English language teachers.* Cambridge University Press.

Burns, A. (2010). *Doing action research in English language teaching: A guide for practitioners.* Routledge.

Burns, A. (2019). Action research in second language teaching: recent developments. *The Language Teacher*, *43*(3), 3–7.

Hutton, M. (1989). Learning from action: a conceptual framework. In S. W. Weil, & I. McGill (Eds.), *Making Sense of Experiential Learning: Diversity in Theory and Practice* (pp. 25–37). UK Society for Research into Higher Education and Open University Press.

Kolb, A., & Kolb, D. (2009). Experiential learning theory: a dynamic, holistic approach to management learning, education and development. In S. J. Armstrong & C. V. Fukami (Eds.), *The SAGE Handbook of Management Learning, Education and Development* (pp. 42–68). SAGE Publications Ltd.

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues*, *2*(4), 34–46.

Malinen, A. (2000). *Towards the essence of adult experiential learning*. University Printing House.

McLeod, A. (2017, Oct 24). *Kolb - learning styles*. https://www.simplypsychology.org/learning-kolb.html

Norma, H. (2010). Experiential learning: theory and challenges. *Christian Education Journal, 7*(1), 107–117.

Paulus, T., & Wise, A. (2019). Looking for insight, transformation, and learning in online talk. *The SAGE Handbook of Online Research Methods* (pp. 340–356). SAGE.

Stringer, E. T., & Aragón, A. O. (2021). *Action research*. SAGE.

**Contact email:** pcwlee@hkmu.edu.hk

# Playing Heritage: Loose Parts and Traditional Emirati Play as Pathways to Culture and Identity in Early Childhood

Alueshima Jennifer Shaapera, Blossom Nursery & Preschool, United Arab Emirates
Nguyilan Shaaji, Emirates National School, United Arab Emirates

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Early years settings commonly approach culture through what can be seen and displayed: national day celebrations, dress-up corners, food festivals, and themed weeks. Such experiences have their place, but they tend to hand culture to children as content rather than as something children live and remake. This paper reports an action research study carried out across five early years settings in the United Arab Emirates, in which educators drew on traditional Emirati play to design open-ended loose parts invitations for children aged three and four. The materials were deliberately ordinary—shells, stones, seeds, sand, fabric, baskets, and recycled objects—chosen both for their resemblance to what Emirati children once played with and for their refusal to dictate a single use. Across successive cycles of planning, enactment, observation and reflection, we generated pedagogical documentation, educator reflective journals and structured observations, and analysed them thematically. Loose parts worked as cultural mediators: shells became pearls, lengths of fabric became majlis spaces, and sand became terrain for storytelling and role play. Values associated with hospitality, generosity, collaboration and community recurred across all five settings, and they surfaced in what children did rather than in what they could say. Children also departed from the traditions that had inspired the invitations, folding in their own preoccupations and inventing variants—behaviour better described as interpretive reproduction than as transmission. We argue that culturally grounded loose parts play offers early years settings a workable alternative to representational approaches, and that it positions young children as participants in cultural life rather than as its audience.

*Keywords:* loose parts, cultural identity, Emirati heritage, play-based pedagogy, early childhood education, action research

iafor
The International Academic Forum
www.iafor.org

## Introduction

Walk into most early years classrooms during a national celebration and culture is easy to spot. It is on the walls, in the costume box, on the snack table. What is harder to see is whether any of it has travelled inward—whether the three-year-old who wore the kandura on National Day carries anything of it into Tuesday morning's block corner. That gap between cultural display and cultural experience is the problem this study set out to work on.

The early years are a period in which children begin to locate themselves in relation to others, and the research consensus is that they do so through participation rather than instruction. Vygotsky (1978) placed cultural tools and social interaction at the centre of development, and Rogoff (2003) extended the argument by treating culture not as a variable acting on children but as the medium in which their development takes place. From this vantage point, a child learns what generosity means by being on both sides of it, not by being told a story about it. Corsaro (1992) pushed the point further: children do not simply absorb the adult world, they appropriate fragments of it and rework them into peer cultures of their own, a process he termed interpretive reproduction.

Curricular practice has been slower to follow. Culture in many settings still arrives as a topic—a week, a display, a visiting artefact—with the result that rich practices are compressed into symbols that sit at some distance from children's ordinary days. Gay (2018) and Ladson-Billings (1995) have both argued that culturally responsive education has to move past the celebratory register and into the daily fabric of pedagogy; Moll et al. (1992) made the complementary case that the knowledge children bring from home is a curricular resource rather than a decorative one. The critique is well established. What is less settled is what it looks like in a nursery classroom on an ordinary morning.

The UAE context sharpens the question in a particular way. Traditional Emirati children's games were inscribed on UNESCO's List of Intangible Cultural Heritage in Need of Urgent Safeguarding in 2011, on the grounds that of roughly two hundred games documented by researchers in the 1990s, only twenty to thirty were still played (UNESCO, 2011). These were games built from what was to hand—sand, shells, stones, sticks, cloth, discarded metal—and organised around gathering, cooperation and improvisation. Baker (2018), working with fifty-two Emirati kindergarten teachers, described the resulting task facing educators as one of enriching a largely imported play pedagogy with Emirati cultural meanings rather than importing a further set of activities. Alhosani (2022) similarly found that while the national kindergarten curriculum framework foregrounds cultural and national identity, its delivery leans on content areas such as social and Islamic studies. Al-Qinneh and Abu-Ayyash (2022) add a further complication from the home side, reporting that a majority of Emirati children in their sample rarely engaged in undirected play at all.

There is, then, a strong policy commitment to heritage, a documented decline in the play practices that once carried it, and a pedagogical tradition of open-ended play that has not often been asked to do cultural work. Loose parts pedagogy seemed to us a plausible bridge. Nicholson (1971) proposed that the inventiveness available in any environment is proportional to the number of variables in it; the materials that filled traditional Emirati play were, in his terms, almost pure variables. The question was whether putting such materials back into the hands of very young children, framed by educators who understood what they had once meant, would produce anything more than pleasant sensory play.

We therefore asked: *How does loose parts pedagogy informed by traditional Emirati play practices support young children's understanding of culture and identity in early childhood settings?*

## Literature Review

### Culture, Identity and Belonging in the Early Years

Culture, in the sense used here, is not a set of artefacts. Hall (1997) treated it as a system of shared meanings through which people make sense of their world and of each other, which places practices and relationships, rather than objects, at its centre. This matters pedagogically: if culture is meaning-making, then a display case cannot carry it, but a shared routine can.

Belonging is the strand of identity work most visible in early years settings, and it has been written into curriculum documents accordingly. The Australian Early Years Learning Framework places belonging alongside being and becoming as the ground of children's wellbeing and learning (AGDE, 2022). Sumsion and Wong (2011) offered a useful corrective to easy readings of that term, mapping belonging across emotional, social, spatial, temporal, cultural and political dimensions and noting that curriculum framings tend to be eloquent about the first few and silent about the last. Their analysis was helpful to us in a practical way: it kept us asking not only whether children felt comfortable, but whose ideas were shaping the play and who was in a position to include or exclude.

Running through this literature is a view of children as competent social actors. James and Prout (2015) argued that childhood is socially constructed and that children are worth studying as agents in their own right rather than as adults in preparation. That commitment has direct methodological consequences, discussed below.

### Loose Parts and Open-Ended Materials

Nicholson's (1971) theory of loose parts is short, polemical and widely cited, and its core claim is deceptively simple: environments rich in movable, combinable, undefined materials generate more inventive behaviour than environments furnished with objects that already know what they are for. Daly and Beloglovsky (2015) translated the idea into early years practice, and Gull et al. (2019), in a scoping review, showed how loosely the term has been applied since—spanning natural materials, manufactured components and recycled objects, with definitional boundaries that shift by author.

The evidence base deserves an honest characterisation. Gibson et al. (2017) reviewed quantitative studies of loose parts interventions and concluded that the volume and quality of evidence were insufficient to establish effects on cognitive, social or emotional development, despite the enthusiasm of practitioners and policymakers. More recent syntheses reach similar conclusions about indoor loose parts play. This is not an argument against the pedagogy; it is an argument for modesty about what any single study, including this one, can claim. Much of the persuasive work in this field is qualitative and concerned with the character of children's engagement rather than with measurable gains, and that is the tradition in which we locate ourselves.

What is largely missing from the loose parts literature is culture. Materials tend to be discussed in terms of affordance, creativity and physical activity, with little attention to the fact that a

shell in Ras Al Khaimah is not the same object as a shell in a Scandinavian forest school. It is that omission the present study addresses.

### Traditional Play as Cultural Practice

Sutton-Smith (1997) warned against tidy accounts of what play is for, and his ambivalence is worth keeping in view when traditional games are enlisted for educational purposes. Games do carry language, social norms and community values across generations, and UNESCO's (2011) nomination file makes exactly that case for the Emirati repertoire, noting that many games were accompanied by songs and lyrical exchanges and were learned by observation rather than instruction. But treating games primarily as vehicles of transmission risks a museum logic in which authenticity is measured by fidelity to a remembered original.

Corsaro (1992) offers the more productive frame. If children routinely appropriate and rework what they are given, then a traditional game reproduced exactly is arguably less culturally alive than one children have altered. Baker (2018) reaches a compatible conclusion from the Abu Dhabi context, arguing for teacher awareness of the cultural persuasions embedded in play rather than for the reinstatement of particular games. This distinction—between reviving practices and reactivating the principles inside them—became the organising decision of our pedagogical design.

## Methodology

### Design

We used cyclical action research, following the planning, acting, observing and reflecting spiral described by Kemmis et al. (2014). The choice was not incidental. The question we were asking is one that only becomes answerable through repeated attempts: a single set of play invitations tells you very little, whereas the third or fourth iteration, redesigned in light of what children actually did with the second, tells you a good deal. Action research also positions educators as practitioner-researchers, which suited a study concerned with what culturally grounded provision requires of adults as much as with what children do.

Each cycle ran as follows: educators planned an invitation grounded in an identified cultural principle; the invitation was set out and left genuinely open; children's engagement was documented; the team reflected collectively; and the next invitation was adjusted. Adjustments were frequently subtractive—removing a prompt, a prop or an adult framing that had been narrowing children's options.

### Settings and Participants

The study was conducted across five early years settings in the United Arab Emirates, involving children aged three to four years and their educators. As is typical of UAE early years provision, the classrooms were culturally and linguistically mixed, with Emirati children learning alongside peers from a wide range of national backgrounds, and staff teams similarly diverse. This is worth stating plainly rather than treating as background: the study was not an exercise in delivering Emirati heritage to Emirati children. It was an exercise in whether the cultural principles carried by local traditional play could become shared ground in classrooms where most children were growing up in the UAE without inheriting those traditions at home.

### Designing the Play Invitations

Educators did not set out to reconstruct historical games. Instead, working from locally documented accounts of traditional play and from staff and family knowledge, they identified the principles those games appeared to run on: hospitality, cooperation, gathering and community, resourcefulness, storytelling, and social participation. Invitations were then designed to make those principles available rather than to require them.

Materials were chosen for their ordinariness and their ambiguity: shells, stones, sand, seeds, fabric pieces, wooden resources, natural baskets and recycled objects. Nothing in the collection announced a use. A basket could be a boat, a majlis floor, a pearl store or a hat. The resemblance to the materials of traditional Emirati childhood was deliberate, but the resemblance did the work through the materials themselves rather than through explanation.

Adults framed and then withdrew. Where a traditional practice was introduced at all, it was introduced as a story or a gathering rather than as a set of rules to be followed.

### Generating Data

Three sources were used. Pedagogical documentation—photographs, children's constructions, learning stories and educator notes—was gathered in the Reggio-influenced sense described by Rinaldi (2006), that is, as material for interpretation and dialogue rather than as a record of achievement. Educator reflective journals tracked observations, uncertainties, design changes and professional shifts, and proved particularly useful for surfacing moments when adults had inadvertently closed play down. Structured observations focused on children's interactions, language, participation patterns and evidence of cultural meaning-making.

### Analysis

Data were analysed thematically, following the reflexive approach set out by Braun and Clarke (2006, 2022). Documentation, journals and observation records were read repeatedly, coded, and organised into candidate themes; themes were then tested against the full data set and against material from each of the five settings in turn. Where a pattern appeared in only one setting it was retained as a note rather than promoted to a theme. Collective interpretation across the educator team, and the comparison of themes across settings, served the trustworthiness of the account in the sense described by Lincoln and Guba (1985), though we make no claim that our reading is the only defensible one.

## Findings and Discussion

### Loose Parts as Cultural Mediators

The clearest finding was also the simplest: the materials did cultural work. Shells became pearls. Stones marked out gathering places. Lengths of fabric were spread and re-spread until they were majlis spaces, and sand became terrain across which stories were told and journeys taken. None of these transformations was suggested. They emerged from children handling objects that were open enough to become something and familiar enough to become something in particular.

This is more than the standard loose parts observation about imaginative flexibility. A shell can become anything; in these settings it repeatedly became a pearl. The materials appeared to carry a local semantic pull, activated by their placement in a setting where pearling, gathering and hospitality were part of the surrounding conversation. This suggests a qualification to how loose parts are usually theorised. Nicholson (1971) framed variables in largely formal terms—number, movability, combinability—and much subsequent work has followed him (Gull et al., 2019). What we observed is better described as cultural affordance: materials whose openness is real but not neutral, weighted by the meanings available in the setting around them. Culture, in other words, was mediated through objects and action, in a manner consistent with Vygotsky's (1978) account of cultural tools, rather than delivered through explanation.

## Culture Enacted Rather Than Recited

Children demonstrated cultural understanding through what they did, not through what they could report. Asked directly about hospitality, a three-year-old has little to say. Observed over a morning, the same child repeatedly invited peers into a constructed gathering space, made room, and handed over materials without being prompted to share.

These actions align closely with values of hospitality, generosity and community participation, and they were among the most consistent patterns across all five settings. We are cautious about the inferential step here. Enacted hospitality in play is not proof that a child has internalised a cultural value, and it would be possible to read the same behaviour simply as sociable three-year-olds being sociable. What can be said with more confidence is that the provision made these particular forms of participation available and that children took them up readily, which is precisely what a representational approach to culture does not do. Hall's (1997) framing is useful: if culture is shared meaning, participation in the practice is not a route to the meaning but part of the meaning itself.

## Collaboration and the Making of Common Ground

Collaborative play emerged in every setting. Children negotiated roles, pooled materials, solved construction problems jointly, and built narratives that outlasted individual contributions. The possessive plural was everywhere in their talk—”our home”, “our place”, “come sit with us”—and its frequency is itself informative. A group of children who have jointly assembled a space out of stones and fabric have made something that belongs to none of them individually.

Read through Sumsion and Wong's (2011) dimensions, what is happening here is spatial and social belonging generated by children rather than conferred by adults. It also raises the political dimension they identify as under-examined. “Come sit with us” is an invitation, but invitations imply a threshold and someone standing at it. Our documentation captured a great deal of inclusion and comparatively little of its opposite, which may reflect the reality of these classrooms, or may reflect what educators found it comfortable to document. We flag it as an open question rather than resolve it.

## Agency and Cultural Authorship

Children did not reproduce traditional practices. They altered them, and the alterations were the most theoretically interesting thing we observed. Constructed gathering spaces acquired functions well outside their traditional referent, becoming community centres of the children's

own devising. Boats stopped being pearling boats and became vehicles for adventure narratives. Shell collections evolved into systems of exchange with rules that had to be explained to newcomers.

Judged against a preservation standard, this is drift. Judged against Corsaro's (1992) account of interpretive reproduction, it is exactly what cultural vitality looks like: children taking material from the adult world, working it over with their peers, and producing something that belongs to their own culture while remaining connected to its source. The finding matters practically for heritage education in the UAE. UNESCO's (2011) safeguarding rationale rests on transmission, and transmission is often operationalised as accurate reproduction. Our observations suggest that a game children have changed may be in better health than one they have merely performed correctly, and that the safeguarding question for early years settings is less how to preserve the games than how to keep the principles inside them in circulation. This is consistent with Baker's (2018) argument for cultural awareness in pedagogy over the reinstatement of specific practices. It also puts pressure on deficit readings of very young children. Three- and four-year-olds here were not receiving culture pending the age at which they might contribute to it. They were contributing already.

### Identity and Belonging as Process

Identity did not appear in our data as an outcome that could be observed arriving. It appeared as something continuously produced in relation to others—in being invited, in having a suggestion adopted, in being the one who knows how the shell exchange works. Opportunities to contribute ideas and shape the direction of play were associated with visible increases in children's ownership of the learning environment.

Given the composition of these classrooms, this has a further implication. Children who had no prior connection to Emirati traditions participated in the same play, on the same terms, and helped author what it became. The cultural material did not function as a marker separating those who owned it from those who did not; it functioned as common ground. For settings serving highly diverse populations, this may be the most useful practical finding in the study, and it echoes Moll et al.'s (1992) point that cultural knowledge becomes educationally powerful when it is put into circulation rather than kept as the property of those it originates with.

## Limitations

Several constraints bound what can be claimed. The study was small, involving five settings and a single age band, and action research of this kind is designed to improve local practice rather than to support generalisation; readers will need to judge transferability against their own contexts. Educators were both practitioners and researchers, which gave us access to interpretation that an external observer could not have reached but also meant that those with a stake in the invitations were the ones deciding what to document. We took cross-setting comparison and collective interpretation as partial safeguards, but they are partial.

The data are also adult-generated. Despite a theoretical commitment to children as social actors (James & Prout, 2015), children's own accounts of what they were making and why were not systematically gathered, and a study that took that commitment fully seriously would find methods for doing so. Finally, we documented what happened during the cycles and cannot speak to what persisted afterwards. Whether the dispositions we observed have any durability is a question this design cannot answer.

## Conclusion and Implications

Loose parts pedagogy informed by traditional Emirati play gave children in five UAE settings a way into culture that display-based approaches do not offer. Ordinary materials—shells, stones, sand, fabric—functioned as cultural mediators, and children used them to enact hospitality, generosity, cooperation and belonging rather than to represent them. They then changed what they had been given, which we read, following Corsaro (1992), as evidence of cultural participation rather than of failed transmission.

For educators, the practical implications are less about materials than about restraint. The invitations that worked were the ones adults framed and then left alone; the recurring correction across cycles was the removal of adult scaffolding that had quietly specified an outcome. Cultural grounding in this model is a matter of what is made available and of how the surrounding conversation is held, not of instruction.

For leaders and policymakers, the study offers a modest case that heritage objectives and play-based pedagogy need not be in tension, and that the two are often kept separate by curricular habit rather than by necessity. Given the documented decline in traditional Emirati children's games (UNESCO, 2011) and the equally documented decline in undirected play at home (Al-Qinneh & Abu-Ayyash, 2022), settings that can address both at once are worth resourcing.

Three directions seem most worth pursuing. Families and elders were absent from this study and are the obvious holders of the knowledge it drew on second-hand. Longitudinal work would test whether anything we observed persists. And the approach should be tried where the heritage in question is not the dominant local one, which would establish whether what we saw depends on cultural congruence with the setting or on the openness of the materials themselves.

## Acknowledgements

We are grateful to the educators, children and families who took part in this study, and to the participating early years settings across the United Arab Emirates for supporting reflective and exploratory practice over successive research cycles. Our colleagues' willingness to have their own pedagogical decisions examined, including the ones that did not work, made the study possible.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that generative AI tools were used to support academic editing, language refinement, the organisation of ideas and the identification of relevant literature during manuscript preparation. All research design, data collection, analysis and interpretation of findings, and all final decisions regarding content, remain the responsibility of the authors. The authors reviewed, verified and approved all content, including all cited sources, prior to submission.

## References

Alhosani, N. (2022). The influence of culture on early childhood education curriculum in the UAE. *ECNU Review of Education, 5*(2), 284–298. https://doi.org/10.1177/20965311221085984

Al-Qinneh, D., & Abu-Ayyash, E. A. S. (2022). The play-based behaviours of Emirati preschool children: Cultural perspective into early childhood education. *Child Care in Practice, 28*(3), 394–410. https://doi.org/10.1080/13575279.2020.1816531

Australian Government Department of Education [AGDE]. (2022). *Belonging, being and becoming: The Early Years Learning Framework for Australia* (V2.0). Australian Government Department of Education for the Ministerial Council.

Baker, F. S. (2018). Shaping pedagogical approaches to learning through play: A pathway to enriching culture and heritage in Abu Dhabi kindergartens. *Early Child Development and Care, 188*(2), 109–125. https://doi.org/10.1080/03004430.2016.1204833

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.

Corsaro, W. A. (1992). Interpretive reproduction in children's peer cultures. *Social Psychology Quarterly, 55*(2), 160–177. https://doi.org/10.2307/2786944

Daly, L., & Beloglovsky, M. (2015). *Loose parts: Inspiring play in young children*. Redleaf Press.

Gay, G. (2018). *Culturally responsive teaching: Theory, research, and practice* (3rd ed.). Teachers College Press.

Gibson, J. L., Cornell, M., & Gill, T. (2017). A systematic review of research into the impact of loose parts play on children's cognitive, social and emotional development. *School Mental Health, 9*(4), 295–309. https://doi.org/10.1007/s12310-017-9220-9

Gull, C., Bogunovich, J., Levenson Goldstein, S., & Rosengarten, T. (2019). Definitions of loose parts in early childhood outdoor classrooms: A scoping review. *The International Journal of Early Childhood Environmental Education, 6*(3), 37–52.

Hall, S. (Ed.). (1997). *Representation: Cultural representations and signifying practices*. SAGE.

James, A., & Prout, A. (Eds.). (2015). *Constructing and reconstructing childhood: Contemporary issues in the sociological study of childhood* (3rd ed.). Routledge.

Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The action research planner: Doing critical participatory action research*. Springer.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465–491. https://doi.org/10.3102/00028312032003465

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice, 31*(2), 132–141. https://doi.org/10.1080/00405849209543534

Nicholson, S. (1971). How not to cheat children: The theory of loose parts. *Landscape Architecture, 62*(1), 30–34.

Rinaldi, C. (2006). *In dialogue with Reggio Emilia: Listening, researching and learning*. Routledge.

Rogoff, B. (2003). *The cultural nature of human development*. Oxford University Press.

Sumsion, J., & Wong, S. (2011). Interrogating 'belonging' in *Belonging, Being and Becoming: The Early Years Learning Framework for Australia*. *Contemporary Issues in Early Childhood, 12*(1), 28–45. https://doi.org/10.2304/ciec.2011.12.1.28

Sutton-Smith, B. (1997). *The ambiguity of play*. Harvard University Press.

UNESCO. (2011). *Children's traditional games in the United Arab Emirates*. List of Intangible Cultural Heritage in Need of Urgent Safeguarding (Decision 6.COM 8.22). United Nations Educational, Scientific and Cultural Organization.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

# Bridging Policy and Practice in India: Implementing IB Inquiry-Based Pedagogy in a Tier-3 CBSE School

Ramya Damarla, Helios Sadineni Chowdaraiah Futuristic School, India

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This study investigates the feasibility of adapting International Baccalaureate (IB) inquiry-based pedagogical methods within urban CBSE middle-school classrooms to bridge the gap between the student-centred, critical-thinking and holistic-assessment vision of India's National Education Policy (NEP) 2020 and National Curriculum Framework for School Education (NCF-SE) 2023 and the traditional examination-driven reality that continues to influence many urban classrooms. Using a mixed-methods research design, data were collected from 20 teachers through structured questionnaires and in-depth interviews, complemented by direct classroom observations and teaching interventions across three Grade 6, 7 and 8 sections involving approximately 100 students in one urban CBSE school. The study examined six skill domains: critical thinking, collaboration, communication, problem-solving, creativity and self-directed learning, alongside student engagement and assessment practices. Findings indicate that teachers strongly recognise the value and policy alignment of inquiry-based approaches, while facing barriers including overloaded syllabi, examination-driven assessment systems, limited professional development opportunities and insufficient institutional support. Teachers receiving structured facilitation support demonstrated improvements in instructional practices, accompanied by higher levels of student engagement and conceptual participation. The study concludes that policy intent alone is insufficient to transform classroom realities. Meaningful change requires targeted teacher capacity building, competency-aligned assessment reform and institutional cultures that actively enable inquiry-based learning to take root and scale within existing urban CBSE frameworks.

*Keywords:* inquiry-based learning, IB pedagogy, CBSE, NEP 2020, experiential learning

iafor
The International Academic Forum
www.iafor.org

## Introduction

India's National Education Policy (NEP) 2020 marks a significant policy shift towards holistic, competency-based and learner-centred school education. It emphasises critical thinking, creativity, problem-solving, conceptual understanding, experiential learning and the development of learners who can apply knowledge meaningfully. The National Curriculum Framework for School Education (NCF-SE) 2023 provides a curricular and pedagogical framework through which these aspirations can be translated into school practice. The policy direction therefore creates a strong rationale for moving beyond instruction dominated by memorisation and recall.

However, policy aspiration and classroom implementation do not automatically move together. Teachers in examination-oriented contexts must work within prescribed syllabi, fixed academic calendars, expectations of curriculum completion and assessment systems that may continue to reward recall. These pressures can reduce the time available for questioning, investigation, discussion and reflection. The resulting gap is not simply a pedagogical problem; it is an implementation problem involving curriculum, assessment, teacher capacity and school culture.

Recent literature reinforces this concern. Rajesh (2026) argues that enquiry-based learning can address the gap between what students learn and their ability to apply knowledge, while identifying systemic barriers to wider adoption. Pandey (2025) reports teacher concerns involving workload, practical training, resources and examination alignment. Karmakar (2025) similarly identifies teacher-capacity, infrastructure and assessment challenges in implementing NEP 2020. These studies provide a useful context for examining inquiry adoption in an urban CBSE setting.

This study explores whether selected pedagogical principles associated with IB inquiry-based learning can be adapted within existing CBSE classrooms. The purpose is not to replace CBSE with IB, but to examine whether inquiry, conceptual understanding, student participation and reflection can be embedded within existing national curricular content. The central concern is therefore how international pedagogical ideas can be contextualised to Indian classrooms while remaining aligned with NEP 2020 and NCF-SE.

## Review of Literature

### Rote Learning and Inquiry Adoption

Rajesh (2026) describes the persistence of rote learning in India as a barrier to transferable learning and argues for wider adoption of enquiry-based approaches. The paper also highlights governance and stakeholder-agency issues. For the present study, this is significant because inquiry requires teachers to make professional decisions about questions, scaffolding and learning experiences rather than merely follow a fixed sequence of information delivery.

Pandey (2025), drawing on qualitative evidence from 120 teachers in Uttarakhand, identifies infrastructure and resource constraints, teacher workload and limited practical training, and misalignment with high-stakes assessment as barriers to inquiry-based learning. Although the present study is located in an urban CBSE school, the assessment and professional-capacity issues are relevant across contexts. Karmakar (2025) further connects NEP 2020 with

constructivist and social-constructivist approaches and identifies teacher capacity and rigid assessment as implementation challenges.

### Experiential and Subject-Based Pedagogical Shifts

Bahadur and Singh (2025) discuss a shift in mathematics education from formula memorisation towards conceptual visualisation and active inquiry. Mondal and Bhat (2026) emphasise practical work, experimentation and experiential learning in science education. Together, these studies support a broader principle: learners require opportunities to investigate, apply and discuss concepts rather than only receive information.

### Teacher Capacity and Assessment

Panda and Lenka (2026) emphasise continuous professional development and the development of teacher capacity under NEP 2020. Banwari (2025) similarly identifies professional learning and educational resources as important for pedagogical innovation. Inquiry-based teaching requires practical competence in questioning, scaffolding, discussion management and formative assessment. Assessment is equally important. Nayak (2024) discusses the movement towards multidimensional assessment and the 360-degree progress card under NEP 2020. If classroom pedagogy promotes reasoning but assessment rewards primarily recall, teachers receive contradictory signals.

The wider inquiry literature also supports structured guidance. Lazonder and Harmsen (2016) found that the effects of inquiry learning vary with the guidance provided. This study therefore treats inquiry as structured exploration rather than unguided discovery.

## Research Questions

1. How feasible is it to adapt IB inquiry-based pedagogical methods within urban CBSE middle-school classrooms?
2. How do inquiry-oriented interventions influence student engagement and the development of critical thinking, collaboration, communication, problem-solving, creativity and self-directed learning?
3. What systemic barriers do teachers experience when transitioning from traditional instructional approaches towards inquiry-based pedagogy, and what support structures are required?

## Methodology

### Research Design

A mixed-methods design was used, combining structured teacher questionnaires, in-depth interviews, direct classroom observations and teaching interventions. The design enabled the study to consider teacher perceptions alongside classroom-level evidence of implementation.

### Participants and Setting

The study was conducted at Helios Sadineni Chowdaraiah Futuristic School in Chilakaluripet, a Tier-3 town in Andhra Pradesh, India. The student component involved approximately 100 learners across three Grade 6, 7 and 8 sections. Twenty teachers participated through structured

questionnaires and in-depth interviews. The study is therefore intended as a contextual investigation of feasibility rather than a statistically generalisable study of all CBSE schools.

## Data Collection and Analysis

**Table 1**
*Summary of the Study Design and Data Sources*

| Component | Scope | Purpose |
|---|---|---|
| Teacher questionnaires | 20 teachers | Perceptions, barriers and perceived outcomes |
| In-depth interviews | 20 teachers | Implementation experiences and support needs |
| Classroom observations | Grades 6–8; ~100 students | Engagement, questioning, collaboration and conceptual participation |
| Teaching interventions | Three middle-school sections | Feasibility of inquiry within existing CBSE content |

Teacher questionnaires examined perceptions of inquiry, student engagement, critical thinking, collaboration, communication, creativity, self-directed learning, curriculum pressure, assessment expectations, professional-development needs and institutional support. Interviews explored the difference between recognising the value of inquiry and implementing it consistently. Classroom observations focused on questioning, participation, peer interaction, teacher questioning, conceptual discussion, collaboration, application and reflection. The available study materials report qualitative and structured findings but do not provide numerical effect sizes or a controlled comparison; consequently, the findings are interpreted as contextual evidence rather than causal estimates.

## Pedagogical Intervention

The intervention followed an adaptation rather than replacement model. Existing CBSE topics were retained while the learning process was redesigned around provocation, questioning, investigation, conceptual clarification, application and reflection. The teacher remained responsible for academic rigour and scaffolding while students received greater opportunities to investigate ideas and articulate understanding.

## Story-Based Inquiry in Civics

One intervention used a story-based approach to introduce concepts concerning forms of government. Rather than beginning only with definitions, the narrative provided a provocation through which students considered power, participation and decision-making. Students then compared systems, discussed the distribution of power and connected their reasoning to formal civic concepts. The sequence moved from experience and questioning towards conceptual understanding.

## Six Skill Domains

The study examined critical thinking, collaboration, communication, problem-solving, creativity and self-directed learning. These were considered alongside engagement because inquiry is expected to change not only what students learn but how they participate in learning. Observation therefore considered whether students asked questions, explained reasoning, interacted with peers and applied ideas.

# Findings

## Teacher Recognition of Inquiry

Teachers strongly recognised the value of inquiry-based learning and its alignment with NEP 2020 and NCF-SE. They associated inquiry with deeper conceptual learning, critical thinking, participation and transferable skills. This suggests that the principal obstacle was not a rejection of learner-centred education, but the conditions under which teachers were expected to implement it.

## Curriculum and Time Pressure

Teachers consistently identified syllabus load and curriculum-completion pressure as major constraints. Inquiry requires time for students to formulate questions, investigate, discuss and reflect. In tightly scheduled academic contexts, teachers may therefore perceive inquiry as difficult to sustain even when they value it.

## Examination-Driven Assessment

Assessment emerged as another significant barrier. Teachers reported tension between inquiry-oriented learning and examination systems that can prioritise predictable recall. If teachers expect students ultimately to be judged through conventional examinations, they may feel pressure to prioritise content coverage and direct instruction. This finding is consistent with the assessment concerns identified by Pandey (2025) and the reform direction discussed by Nayak (2024).

## Professional Development and Facilitation

Teachers who received structured facilitation support demonstrated improvements in instructional practices. Practical support in framing questions, sequencing activities and connecting inquiry to curricular objectives made implementation more manageable. This indicates that professional development should move beyond explaining the theory of inquiry and provide teachers with opportunities to design, implement and refine actual lessons.

## Student Engagement and Conceptual Participation

Classroom observations indicated higher levels of student engagement and conceptual participation during structured inquiry-oriented activities. Students had more opportunities to ask questions, discuss ideas, compare perspectives and articulate reasoning. These observations should not be interpreted as proof of a causal effect because the study did not employ a controlled experimental design or report effect-size statistics.

## The Six Skill Domains in Practice

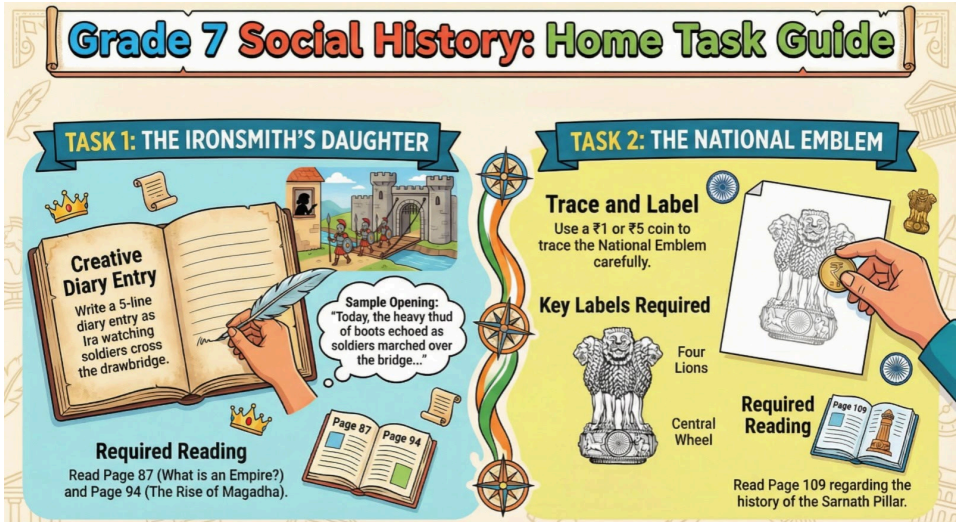
The six domains provided a useful lens for interpreting classroom activity. Critical thinking was reflected in comparison and justification; collaboration in peer discussion; communication in explanation; problem-solving in applying concepts; creativity in generating responses; and self-directed learning in students' increasing responsibility for questions and reflection. The findings therefore suggest that inquiry can create a common classroom structure through which several competencies are practised simultaneously.

## From Classroom Inquiry to Home-Learning Inquiry

A particularly visible feature of the implementation at Helios School was that inquiry was not confined to the classroom period. The same inquiry logic was extended into home-learning tasks so that students could observe, question, investigate, apply, communicate and act beyond the textbook. The tasks were deliberately designed as short, visually structured prompts rather than as conventional instructions to read a chapter and answer questions. This is important in a Tier-3 town context because the approach used familiar surroundings, family conversations, local civic life, weather observations, everyday environmental choices and accessible materials as learning resources.

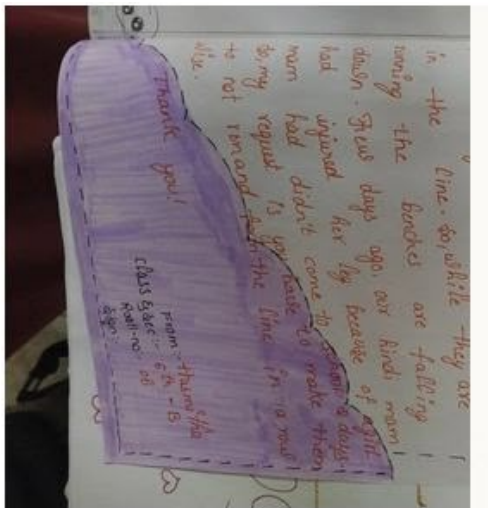
The home-learning materials also show a form of digital and visual task design. The Grade 6 social task sheet, for example, asks learners to conduct a local-needs survey, identify a local elected representative and ask family members about participation in a Gram Sabha or community meeting. The Grade 7 geography tasks ask learners to keep a 24-hour weather and wardrobe diary, construct a simple rain gauge, record rainfall and observe wind. The Grade 7 social-history task asks learners to write a creative diary entry from a historical perspective and trace and label the national emblem using a coin. These are not simply decorated worksheets: the visual layout functions as a scaffold that makes the inquiry question, procedure, evidence to collect and expected product visible to the learner (Figure 1).

**Figure 1**
*Digitally Designed, Illustrated Home-Learning Tasks Used to Extend Inquiry Beyond the Classroom*




The environmental task provides another example. Students are first prompted to think about the impact of oceans and then distinguish positive and negative impacts before proposing one simple action. The "Planet Protector Pledge" converts that reflection into a personal commitment, such as turning off the tap while brushing or reducing plastic-bag use. A handwritten student response to a real-life complaint about an unsafe or poorly maintained public space provides a further example of students moving from observation and concern to written communication and a possible civic response (Figure 2). The significance of these tasks is therefore pedagogical: the assignment itself becomes part of the inquiry cycle rather than a post-lesson recall exercise.

**Figure 2**
*Student-Facing Environmental Prompt and Handwritten Evidence of Real-World Written Response*



The implementation can be interpreted through the inquiry-cycle sequence used in the study: tuning in, questioning, finding out or investigating, sorting and organising information, going further through application, making conclusions and taking action. Not every individual task contains every phase, and the evidence should not be interpreted as a rigid one-to-one mapping. Rather, the collection of classroom and home-learning experiences shows repeated movement between these phases. Students tune in through provocations and everyday situations; formulate or respond to questions; gather information through observation, family discussion, reading and simple measurement; organise evidence; apply concepts to local and real-life contexts; communicate conclusions; and, in selected tasks, propose or undertake action. (Murdoch, 2015).

**Table 2**
*Inquiry-Cycle Evidence Across Classroom and Home Learning*

| Inquiry phase | Practical evidence from Helios School implementation | Competencies made visible |
|---|---|---|
| Tuning in | Environmental pledge; local government and community prompts; historical diary provocation; weather observation | engagement, prior knowledge, questioning |
| Questioning | Local-needs survey, voter check, impact-of-oceans prompt, teacher-guided inquiry questions | critical thinking, communication |
| Finding out / investigating | Family/community conversations, required reading, weather diary, rainfall measurement, wind observation | self-directed learning, problem-solving |
| Sorting and organising | Recording information, distinguishing positive/negative impacts, tracing and labelling, comparing observations | critical thinking, conceptual understanding |
| Going further / applying | Connecting concepts to local government, environmental behaviour, weather processes and historical perspectives | application, creativity, transfer |
| Making conclusions | Short explanatory paragraphs, diary entries, reflections and written responses | communication, reasoning, reflection |
| Taking action | Planet Protector Pledge; proposing simple family actions; communicating a local concern | agency, responsibility, civic participation |

## Discussion

The findings indicate that inquiry adoption in CBSE classrooms is primarily an implementation challenge. Teachers recognise its value, yet curriculum pacing, examination expectations and limited professional support make conventional instruction easier to sustain. This aligns with Rajesh (2026), Pandey (2025) and Karmakar (2025), while also supporting the broader direction of NEP 2020 and NCF-SE.

### IB Principles Within CBSE

A central implication is that curriculum adoption and pedagogical adaptation should be distinguished. The study does not suggest that CBSE schools should become IB schools. Instead, selected IB principles—questioning, conceptual understanding, reflection, student participation and structured inquiry—can be adapted within CBSE content. This creates a potentially more feasible route to pedagogical change because schools can retain their curricular obligations while changing how learning is experienced.

### Teacher as Facilitator

Inquiry does not reduce the importance of the teacher. Effective inquiry requires teachers to design meaningful questions, activate prior knowledge, anticipate misconceptions, provide scaffolding, manage discussion and decide when explicit teaching is necessary. The teacher's role changes from primarily transmitting information to designing and facilitating learning while maintaining academic rigour.

### Assessment Alignment

Assessment must support the same competencies that pedagogy seeks to develop. If students are taught to analyse, discuss and apply knowledge but are assessed mainly through recall, a structural contradiction remains. Formative questioning, performance tasks, portfolios, reflection and application-based questions can complement conventional assessment and make competency development more visible.

### Teacher Development and School Culture

Professional development should be sustained and classroom-based. A practical cycle could involve understanding inquiry, observing examples, co-planning a lesson, implementing it, reflecting on student responses, receiving peer feedback and refining the lesson. School leaders also need to provide planning time and create conditions in which teachers can experiment, reflect and improve.

## Implications for Practice

Teachers can begin with small curriculum-connected changes rather than redesigning every lesson. A familiar topic can be reframed around an authentic question, followed by structured investigation, teacher clarification, application and reflection. School leaders can support this through collaborative planning, peer observation and assessment practices that value reasoning. Teacher education providers can strengthen preparation by making inquiry training practical and classroom-based. Policymakers should continue aligning curriculum, teacher

development and assessment so that policy expectations are not undermined by conflicting institutional incentives.

## Recommendations

1. Develop sustained, practical professional development focused on inquiry-based lesson design and implementation.
2. Provide teachers with adaptable inquiry questions, scaffolds, activities and assessment rubrics connected to existing CBSE content.
3. Strengthen assessment practices so that reasoning, application, reflection and other competencies are assessed alongside knowledge.
4. Protect collaborative planning time for teachers to design and refine inquiry lessons.
5. Encourage school leaders to support pedagogical experimentation and peer learning.
6. Adapt international pedagogical approaches to Indian curricular and cultural contexts rather than importing complete curricular models.
7. Conduct larger, multi-school and longitudinal studies to examine sustainability and student outcomes.

## Limitations

The study was conducted in one urban CBSE school and therefore should not be statistically generalised to all CBSE schools. The teacher sample comprised 20 participants and the student component involved approximately 100 learners across Grades 6–8. The study used observations and interventions rather than a randomised controlled design. The available study materials do not provide numerical effect sizes, pre/post-test scores or detailed instrument reliability statistics. Accordingly, claims about improved engagement and instructional practice are interpreted as contextual evidence rather than causal estimates.

## Conclusion

Ultimately, bridging policy and practice requires coherence between curriculum, pedagogy, assessment and teacher capacity. NEP 2020 and NCF-SE provide the policy direction, while inquiry-based pedagogy offers a practical pathway for developing critical thinking, collaboration, communication, problem-solving, creativity and self-directed learning. The experience documented in this study suggests that meaningful change is possible when inquiry is structured through a clear cycle, embedded in everyday learning, connected to authentic responsibilities and supported by continuous teacher capacity building and assessment reform.

The study therefore recommends that government agencies, local educational bodies and school leadership provide sustained and periodic capacity-building opportunities for teachers and facilitators. Professional learning should be practical, contextualised and followed by opportunities for reflection, classroom application and feedback. Periodic review and adjustment of curriculum, assignments, assessment practices and teacher-support mechanisms should accompany the implementation so that inquiry is not introduced as a one-time initiative but developed as a sustainable school practice.

A major barrier identified through the implementation was a knowledge and capacity gap. Teachers and facilitators may value inquiry-based learning but still require practical knowledge of how to design inquiry questions, scaffold student thinking, manage collaborative learning, assess competencies and move through an inquiry cycle effectively. Teacher and facilitator

capacity building should therefore be treated as a central component of implementation rather than an optional professional-development activity.

The findings suggest that assessment must be considered alongside pedagogy when inquiry-based learning is implemented. Periodic examinations and conventional assessment practices should not operate in isolation from the competencies that inquiry seeks to develop. A more balanced approach should combine ongoing formative assessment, daily learning evidence, reflection, application and competency-oriented tasks with required summative assessment. Such alignment can reduce the tension between inquiry-based classroom practice and examination expectations.

The intervention also sought to connect learning with real-life responsibilities and the Sustainable Development Goals (SDGs), allowing students to understand that learning extends beyond academic tasks. Assignments and learning checks were integrated into students' day-to-day learning rather than being concentrated only in periodic assessment points. This approach enabled more continuous observation of students' understanding, participation and skill development and provided opportunities for timely feedback and adjustment.

The implementation followed an inquiry cycle beginning with tuning in and moving through questioning, investigation, collaboration, application and reflection. This cycle provided students with repeated opportunities to engage with concepts rather than simply receive information. Across the implementation, the researcher observed an approximately 60% improvement in students' critical-thinking performance. This figure is reported as an observation from the implementation and should be interpreted in relation to the study's context and available assessment evidence rather than as a generalisable causal effect.

This study examined the feasibility of adapting IB inquiry-based pedagogy within urban CBSE middle-school classrooms in alignment with NEP 2020 and NCF-SE. Based on the classroom implementation and observations, the study indicates that inquiry-based learning can be meaningfully embedded within the CBSE context when it is deliberately structured, supported and connected to the existing curriculum.

## Acknowledgements

The author gratefully acknowledges the participating teachers and students whose classroom experiences and perspectives contributed to this study, as well as the school leadership and academic community that supported classroom observations and teaching interventions. The author also acknowledges the European Conference on Education (ECE2026) and the International Academic Forum for providing an academic platform for presentation and scholarly discussion.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

Generative AI and AI-assisted technologies were used during manuscript preparation to support language organisation, readability and editorial formatting. The ideas, procedures, findings, analyses, and discussion are originally written and derived from careful and systemic conduct of the research. AI-assisted tools were not used to fabricate participant responses, research observations or empirical data. Grammarly and ChatGPT were used to support

language organisation, readability and editorial refinement, while NotebookLM was used to support the organisation and review of research documents.

## References

Bahadur, T., & Singh, H. K. (2025). Pedagogical shifts in mathematics education: The role of NEP 2020 in shaping curriculum practices. *Journal of Research Innovation and Implications in Education, 9*(4), 110–115.

Banwari, Dr. (2025). Implementation of NEP 2020: Challenges and opportunities in pedagogical innovation. *Scholarly Research Journal for Humanity Science & English Language, 13*(67), 88–94.

Karmakar, T. K. (2025). A theoretical analysis of the National Education Policy 2020 of India and its implementation. *International Journal of Creative Research Thoughts, 13*(7), e91–e94.

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research, 86*(3), 681–718. https://doi.org/10.3102/0034654315627366

Ministry of Education, Government of India. (2020). *National Education Policy 2020*. Government of India.

Mondal, A. K., & Bhat, S. C. (2026). Enhancing science education through practical work: A focus on teacher education in the context of NEP-2020. *Bharati International Journal of Multidisciplinary Research & Development, 4*(5), 39–44.

Murdoch, K. (2015). *The power of inquiry: Teaching and learning with curiosity, creativity, and purpose in the contemporary classroom*. Seastar Education.

National Council of Educational Research and Training. (2023). *National Curriculum Framework for School Education 2023*. NCERT.

Nayak, B. (2024). Latest development in educational assessment with reference to National Education Policy 2020. *International Journal of Creative Research Thoughts, 12*(5), f754–f758.

Panda, S., & Lenka, S. K. (2026). Transforming teacher education in India through NEP-2020: Integrating cognition, creativity and global Indian perspectives. *International Journal of Innovative Research in Social Sciences, 10*(7).

Pandey, S. (2025). Navigating pedagogical shifts: Challenges faced by teachers in implementing inquiry-based learning in rural primary & upper primary schools of Uttarakhand. *International Journal of Research Culture Society, 9*(7), 126–134.

Rajesh, S. (2026). *Debunking rote learning: The case for enquiry-based learning in India* (Working Paper No. 344). Centre for Civil Society.

**Contact email:** ramyad0323@gmail.com

# Appendices

## Appendix A

### *Teacher Questionnaire–Key Domains*

The questionnaire examined teacher understanding of inquiry-based learning; perceived benefits for engagement and higher-order thinking; confidence in implementation; curriculum constraints; assessment pressures; professional-development needs; institutional support; and the six skill domains of critical thinking, collaboration, communication, problem-solving, creativity and self-directed learning.

### *Classroom Observation Focus*

Observation focused on student questioning, participation, peer interaction, teacher questioning, opportunities for student choice, conceptual discussion, collaborative learning, application, reflection and teacher scaffolding.

## Appendix B

### *Additional Classroom and Home-Learning Evidence*

These materials supplement the evidence discussed in the manuscript. They include digitally designed home-learning task sheets, an environmental action pledge, and a student handwritten response. The materials are presented as evidence of inquiry-oriented task design and enactment rather than as decorative worksheet examples.

**Figure 3**
*Digitally Designed and Handwritten Home-Learning Evidence Used to Extend Inquiry Beyond the Classroom*

### *Previously Documented Classroom Evidence*

The following composite retains the classroom photographs supplied for the earlier manuscript and shows additional evidence of questioning, civic participation, collaboration, application and reflection.

**Figure 4**

*Previously Documented Classroom Evidence: Selected Activities*

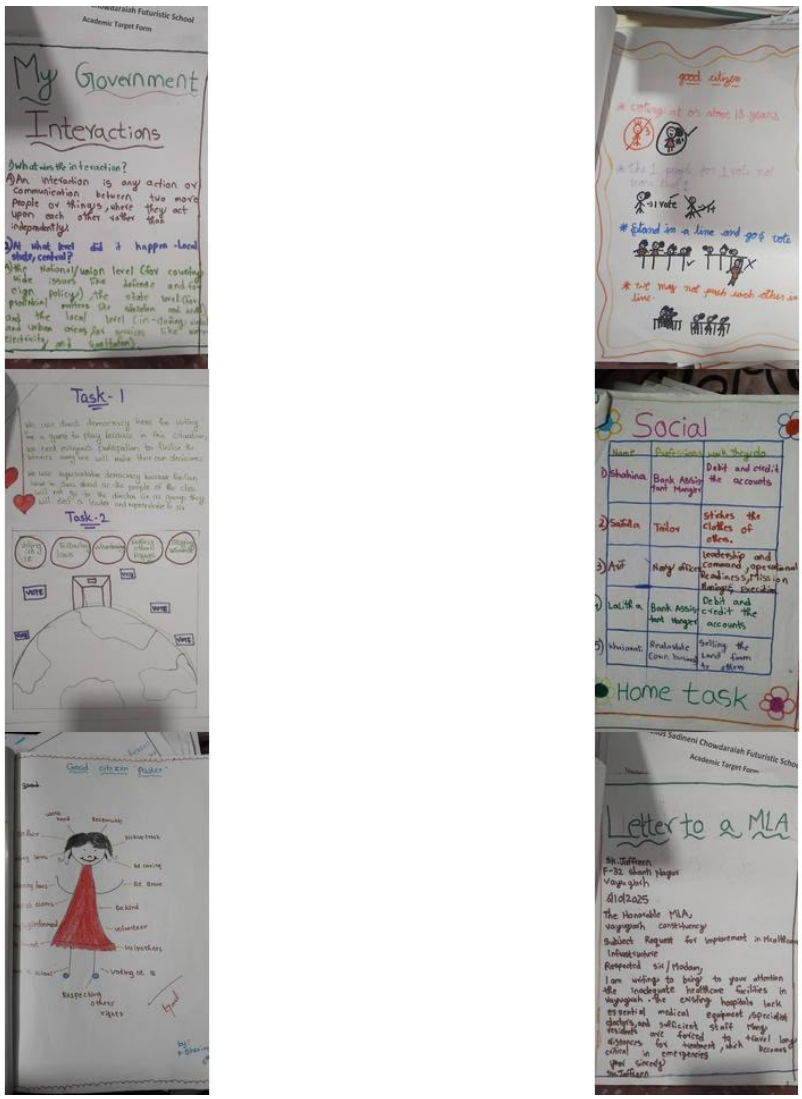


**Figure 5**

*Previously Documented Classroom Evidence: Additional Selected Activities*

# Type One Activations in Suggestopedia: Scaffolding, Suggestion, and the Activation of Hidden Reserves

Christina Johnstone, Restad Gård Utbildning, Sweden

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Type One Activations have become a central feature of contemporary Nordic suggestopedic practice, yet they have received little theoretical attention in the academic literature. This paper develops a theoretical framework for understanding Type One Activations by bringing together Lozanov's theory of Suggestopedia with sociocultural perspectives on scaffolding, mediation and the zone of proximal development. Rather than viewing Type One Activations as individual classroom activities, the paper argues that they are intentionally designed learning environments that support participation while engaging learners across both conscious and paraconscious levels of experience. The analysis examines how dual-plane communication, the activation of hidden reserves, distributed scaffolding and mediated participation interact within these learning environments. It further introduces the concept of families of Type One Activations as an analytical framework for describing recurring patterns of pedagogical design without reducing practice to fixed techniques. Examples drawn from contemporary Nordic suggestopedic practice illustrate how support may be distributed through resources such as the textbook, cards and the organisation of classroom space while preserving the same underlying pedagogical principles. The paper concludes that the educational significance of Type One Activations lies not in the activities themselves but in the relationships they establish within the learning environment. By shifting attention from classroom techniques to the intentional design of environments for supported participation, the proposed framework offers a theoretical language for analysing contemporary suggestopedic practice and contributes to broader discussions of scaffolding, mediation, participation and learning design.

*Keywords:* Suggestopedia, Type One Activations, scaffolding, mediated learning, learning environments

iafor
The International Academic Forum
www.iafor.org

## Introduction

Although Suggestopedia has continued to develop through classroom practice since its original formulation by Georgi Lozanov, some of its most central pedagogical developments have received little theoretical attention. In the Nordic countries, one such development is the concept of Type One activations, a pedagogical distinction that has emerged from classroom practice rather than from Lozanov's own terminology. Despite their central role in organising participation and learning in contemporary suggestopedic classrooms, Type One activations remain largely under-theorised in the academic literature. This absence has limited theoretical understanding of how supported participation, suggestive learning environments and the activation of learners' hidden reserves interact within contemporary suggestopedic practice.

Their limited presence in theoretical discussion stands in marked contrast to their central role in contemporary suggestopedic practice. Rather than representing occasional classroom activities, Type One activations provide the principal pedagogical framework through which learners encounter, explore, and gradually appropriate new material (Johnstone, 2026, p. 123). As the primary form of supported participation within the suggestopedic learning cycle, they constitute a fundamental yet largely unexamined component of contemporary suggestopedic pedagogy.

This paper develops a theoretical account of Type One activations by situating them within the broader framework of Suggestology and Suggestopedia and by relating them to contemporary theories of scaffolding and mediated learning. It argues that Type One activations are not simply classroom activities but carefully designed pedagogical environments that operate simultaneously across conscious and paraconscious levels of experience. Through this dual operation, they support both the desuggestive and suggestive processes described by Lozanov, creating conditions in which learners can gradually access what he termed their hidden reserves.

To illustrate how this pedagogical logic is realised in practice, the paper also introduces the notion of families of Type One activations. These families do not represent fixed categories of activities but recurring patterns in pedagogical design. As an analytical framework, they make it possible to describe how common pedagogical principles are expressed through different forms while preserving the flexibility, creativity and responsiveness that characterise suggestopedic practice. The paper first develops the theoretical foundations of Type One activations before illustrating this framework through recurring families of pedagogical design.

## Theoretical Framework

The theoretical framework developed in this paper brings together complementary perspectives from Suggestopedia and sociocultural theories of learning. It draws on Lozanov's theory of Suggestopedia, particularly his concepts of dual-plane communication and hidden reserves, together with Vygotsky's concept of the zone of proximal development and later work on scaffolding and mediation. Together, these perspectives provide a conceptual framework for understanding Type One activations as intentionally designed learning environments that support participation while creating conditions for learning beyond immediate conscious performance.

## Suggestopedia and Dual-Plane Learning

The first perspective is Lozanov's understanding of Suggestopedia as a form of learning that operates across multiple levels of awareness. Central to this perspective is the concept of dual-plane communication, which proposes that classroom interaction unfolds simultaneously on two interconnected levels. On one level, learners engage consciously with language, structure and meaning. On the other, a continuous field of peripheral, emotional, aesthetic and symbolic influences shapes perception, memory and expectations beyond immediate awareness (Lozanov, 2009, p. 103). Understanding this dual operation is central to the present analysis, as it makes it possible to view Type One activations not simply as classroom activities but as intentionally designed learning environments that engage both planes simultaneously.

The second plane is expressed through elements such as intonation, rhythm, gesture, spatial organisation and aesthetic framing. Rather than functioning as separate techniques, these elements work together to create a suggestive learning environment that shapes learners' expectations, reduces tension and strengthens learners' sense of possibility. Suggestion therefore operates not as an isolated method but as an ongoing process embedded in the total learning environment. From this perspective, learning is influenced not only by what is consciously taught but also by how the educational environment is experienced as a whole.

Within this framework, learning is understood as both conscious and paraconscious. Learners actively process certain aspects of the material while simultaneously absorbing a broader range of linguistic, emotional and contextual input through peripheral perception (Lozanov, 2005, pp. 50–52). Rather than depending solely on deliberate attention, learning emerges through the continuous interaction of these two planes, allowing patterns, meanings and associations to accumulate over time. This understanding illuminates the dual-plane design of Type One activations, which engage learners simultaneously through conscious participation and the wider suggestive field of the learning environment.

## Hidden Reserves and the Desuggestive Process

A second key concept in Lozanov's theory is that of hidden reserves, referring to latent capacities for learning, memory, creativity and personal development that remain only partially realised within conventional educational settings (Lozanov, 2005, p. 62). According to Lozanov, these capacities are not absent but constrained by socially and psychologically constructed limitations, including negative expectations, previous educational experiences and restricted beliefs about one's own ability to learn. The educational challenge is therefore not to create new capacities but to establish conditions in which existing potential can gradually emerge.

Suggestopedia seeks to establish these conditions through the complementary processes of desuggestion and suggestion. The desuggestive process reduces tension, self-doubt and limiting beliefs that constrain learners' participation, while the suggestive process cultivates trust, confidence and a sense of possibility. Together, these processes create an educational environment in which learners are more willing to participate, take risks and gradually express capacities that might otherwise remain dormant. From this perspective, the activation of hidden reserves is not the result of a single method or technique but of the cumulative influence of the learning environment as a whole (Lozanov, 2005, pp. 40–41).

As mentioned above, these processes are mediated through the learning environment as a whole. Language, tone of voice, rhythm, gesture, visual design, spatial organisation and social interaction all contribute to shaping how learners perceive both the learning task and their own capacity to engage with it. Learning thus becomes a process of gradual unfolding, in which existing potential is brought into expression through participation in a carefully designed educational environment. This perspective helps explain the pedagogical logic of Type One activations, whose effectiveness lies not in individual activities but in the orchestration of conditions that support learners' confidence, participation and development.

**Scaffolding, Mediation and the Zone of Proximal Development**

Sociocultural theories of learning provide a complementary perspective, particularly through the concepts of scaffolding, mediation and the zone of proximal development. These concepts offer a contemporary theoretical lens for understanding how support is organised within Type One activations. Scaffolding, introduced by Wood et al. (1976), refers to the structured assistance that enables learners to accomplish tasks beyond their current level of independent competence. As learners develop greater confidence and competence, this support is gradually reduced, allowing responsibility to shift towards independent participation.

From a sociocultural perspective, scaffolding can be further understood through Lev Vygotsky's concept of the zone of proximal development (ZPD). The ZPD describes the distance between what learners can accomplish independently and what they can achieve with guidance, collaboration or appropriate support (Vygotsky, 1978). Learning is understood to occur within this space as learners participate in activities that extend beyond their current level of independent performance. Rather than waiting until competence has fully developed, support enables learners to participate successfully while new knowledge and skills are still emerging.

The support provided through scaffolding is realised through processes of mediation. At a macro level, mediation is embedded in the design of tasks, materials and sequences of activity. At a micro level, it emerges through interaction, as teachers and learners jointly construct understanding through guidance, feedback, modelling and shared attention. Learning is also mediated through cultural artefacts, including language, texts, visual supports and physical objects, as well as through the organisation of the learning environment, all of which shape how learners perceive, interpret and engage with new experiences (Hammond & Gibbons, 2005).

These perspectives bring Suggestopedia and sociocultural theory into dialogue. They provide a conceptual understanding of Type One activations as a form of pedagogical design in which suggestion, mediation and scaffolded participation interact to create conditions for learning beyond immediate independent performance. Rather than viewing support primarily as assistance provided by the teacher, they highlight how participation is distributed across the design of the learning environment itself. Within this framework, Type One activations can be understood as carefully designed learning environments situated within the learner's zone of proximal development. Through mediating resources embedded within pedagogical design, they enable learners to participate meaningfully before independent performance is expected. As participation becomes increasingly confident and autonomous, these supports gradually prepare learners for more independent forms of activation.

## Type One Activations as Dual-Plane Pedagogical Design

Type One activations may appear as engaging or playful classroom activities. This surface impression captures an aspect of their form, yet it does not fully explain their pedagogical function. Drawing on the theoretical perspectives presented above, Type One activations can instead be understood as carefully orchestrated learning environments in which dual-plane communication, the activation of hidden reserves and scaffolded participation are intentionally integrated. Their significance therefore lies not in the individual activity itself but in the pedagogical design that supports participation, confidence and the gradual development of independent performance.

On the conscious plane, Type One activations provide learners with clarity, structure and continuous support for participation. Instructions are explicit, goals are transparent and tasks are carefully scaffolded so that learners can engage confidently within a shared framework. The level of challenge is intentionally calibrated to lie within the learner's zone of proximal development, allowing learners to experience success while new knowledge and skills are still emerging. Language, visual supports and other mediating resources remain available throughout the activity, enabling learners to focus on communication and meaning rather than on overcoming unnecessary barriers to participation.

At the same time, Type One activations operate on the paraconscious plane. They immerse learners in a richer field of linguistic, emotional and aesthetic input than is immediately required for successful task completion, thereby engaging peripheral perception. Vocabulary, grammatical patterns and communicative structures recur with purposeful variation, allowing recognition, familiarity and internalisation to develop without continuous explicit analysis. Tone of voice, rhythm, aesthetic framing and elements of play contribute to a suggestive learning environment that nurtures confidence, reduces tension and maintains continuity with the earlier phases of the suggestopedic cycle. Rather than functioning as background decoration, these elements constitute an integral part of the pedagogical design through which learning extends beyond immediate conscious attention.

The interaction of these two planes enables previously encountered material to be brought into active use. Content encountered during earlier phases of the suggestopedic cycle can re-emerge naturally through participation in carefully structured activities. Learning therefore unfolds gradually and without unnecessary pressure, as supported participation allows confidence, familiarity and competence to develop together. Rather than separating input from use, Type One activations create a continuous process in which participation becomes the bridge from understanding to independent performance.

## Type One Activations as the Organising Principle of Suggestopedic Learning

Within the suggestopedic learning cycle, Type One activations constitute the organising principle through which learning is structured during the activation phase. Lozanov referred to this phase as elaboration, emphasising the gradual development of knowledge through active engagement (Lozanov, 2005, pp. 96–98). Contemporary Nordic suggestopedic practice has extended this understanding by distinguishing between supported and less supported forms of activation. From this perspective, Type One activations represent the supported form of elaboration, creating the pedagogical conditions through which learners progressively move from supported participation towards independent performance.

## Families of Type One Activations

Type One activations take many forms, yet they are not random. Across contemporary suggestopedic practice, recurring patterns emerge in the ways participation, support and meaning are organised within the learning environment. It is therefore useful to speak of families of Type One activations. These families do not represent fixed categories or separate techniques but recurring forms of pedagogical design through which the same underlying principles are expressed in different ways.

The families presented below illustrate different ways in which pedagogical support can be embedded within the learning environment. In the examples that follow, support is distributed through language (the textbook), materials (cards) and space (the room), while the underlying pedagogical logic of scaffolded participation and suggestive design remains constant. They demonstrate how Type One activations create learning environments in which participation precedes independent performance.

Although the families differ in their outward form, they share a common pedagogical logic. Type One activations are designed so that all learners can participate successfully. They invite imagination and creativity, remain anchored in a shared text, dialogue or story world and sustain continuous communication throughout the activity. Together, these characteristics create conditions in which participation remains both accessible and meaningful while suggestion operates through repetition, variation, rhythm, aesthetics and social interaction.

Within contemporary Nordic suggestopedic practice, six broad families can be identified: those based on the textbook or a written dialogue, the room, the body, cards, objects and artefacts and game boards (Johnstone, 2026, p. 126). The present paper focuses on three of these families: the textbook, cards and the room itself as pedagogical resources. These examples illustrate how different pedagogical resources distribute support in complementary ways while preserving the same underlying principles of scaffolded participation, dual-plane learning and continuous communication.

### The Textbook as the Primary Mediating Resource

Type One activations built around the textbook transform written language into shared experience. Sentences, dialogues and scenes become starting points for interaction rather than objects of isolated analysis. Vocabulary, grammatical patterns and communicative expressions remain embedded in the narrative or thematic context, allowing learners to encounter language as meaningful communication rather than as disconnected linguistic elements. The textbook therefore functions as a mediating resource, creating coherence, continuity and rich opportunities for scaffolded participation.

The long choral reading illustrates this pedagogical logic particularly clearly. Conducted on the second day, it builds upon the active and passive concert sessions by enabling learners to bring the text into their own voices for the first time in a highly supportive environment. The teacher reads each line and the group responds together in chorus, allowing every learner to participate successfully without the demands of independent production (Johnstone, 2026, p. 128). At the same time, rhythm, intonation, phrasing and expression become embodied through collective reading, while learners' familiarity with the linguistic patterns of the text continues to develop through both conscious participation and peripheral perception. Rather than serving merely as

repetition, the long choral reading establishes a shared linguistic foundation upon which subsequent Type One activations build.

Short choral readings perform a complementary function throughout the learning cycle. They frequently open or conclude later activations, reconnecting learners with the text and continually renewing the relationship between the global whole and the particular activity. Together, these forms of textbook-based activation demonstrate how a single mediating resource can support participation across both conscious and paraconscious planes while maintaining coherence throughout the learning process.

## Cards as Distributed Scaffolding

Cards represent one of the most versatile mediating resources within Type One activations. They may be used independently or combined with other families, but their pedagogical significance lies in their ability to distribute support directly within the learning environment. Rather than relying solely on teacher guidance, well-designed cards embed linguistic, visual and conceptual support in the cards themselves, allowing learners to participate confidently while maintaining the flow of communication.

This makes the card family a particularly clear illustration of Type One activations as designed learning environments. A well-designed card never presents learners with an empty communicative space. Instead, it provides language, images, prompts, dialogue frames or other mediating resources that make successful participation immediately accessible. At the same time, additional vocabulary, grammatical patterns, symbols or visual details may be encountered through repeated exposure without being explicitly required for the task. In this way, the card simultaneously supports conscious participation and peripheral learning, illustrating the dual-plane organisation of suggestopedic pedagogy.

Cards also sustain continuous interaction. They circulate between learners, change hands, invite questions and generate repeated exchanges in which communication remains central while the scaffolding recedes into the background. Because much of the support is embedded in the material, learners can participate actively without the pressure of unsupported language production.

A domino-based activation illustrates this principle particularly clearly. Learners work with illustrated cards representing dishes or other elements drawn from the shared dialogue. Following the familiar logic of domino matching, each placement naturally gives rise to a short communicative exchange supported by simple dialogue frames. The structure remains predictable, while each interaction introduces new combinations of vocabulary, images and meanings (Johnstone, 2026, p. 143).

Within this format, repetition emerges naturally rather than mechanically. Learners encounter the same language through multiple speakers, contexts and combinations, allowing familiarity and confidence to develop gradually. Visual cues, narrative references and structured language frames work together to support both comprehension and expression, while the familiar game format reduces unnecessary cognitive demands and allows attention to remain focused on communication.

The card family therefore illustrates how support can be distributed throughout the learning environment rather than residing solely with the teacher. Scaffolding remains continuously

available while interaction stays dynamic, enabling learners to recognise patterns, participate successfully and gradually develop independence within an environment in which support is ever-present but rarely intrusive.

### The Room as Spatial Scaffolding

A third family of Type One activations distributes support through the organisation of physical space. In these activations, the classroom is not simply the setting in which learning takes place but an active component of the pedagogical design. Posters, visual displays, objects, furniture, movement pathways and the symbolic organisation of space all contribute to the learning environment. Rather than functioning as decoration, these elements function as mediating resources that support participation across both conscious and paraconscious planes.

The room supports learning in complementary ways. First, it provides a stable field of visual, linguistic and symbolic resources that learners encounter repeatedly, often through peripheral perception. Posters, maps, illustrations and the overall aesthetic organisation of the room contribute to orientation, familiarity and recall while helping establish the suggestive atmosphere characteristic of Suggestopedia. Second, the room functions as a form of spatial scaffolding in which movement through the environment becomes part of how language and meaning are organised, experienced and remembered.

The activity *Four Corners of Time* illustrates this principle particularly clearly. Different areas of the room are associated with different tense forms, and learners move physically in response to spoken sentences. Grammar is therefore experienced not only as verbal explanation but also as spatial organisation. Learners listen, move, formulate responses and recognise recurring patterns within an environment where the room itself continually supports participation (Johnstone, 2026, p. 135). The physical environment thus becomes a mediating resource through which abstract grammatical relationships are transformed into shared experience.

This family also demonstrates that support extends beyond linguistic scaffolding. The arrangement of space, the aesthetic qualities of the room and the emotional atmosphere all influence how learners experience participation. A welcoming and meaningful environment encourages confidence, exploration and interaction while contributing to the suggestive field through which learning extends beyond immediate conscious attention. In this way, the room itself becomes an integral part of the pedagogical design that characterises Type One activations.

## Families as an Analytical Framework

Understanding Type One activations in terms of families makes it possible to describe their underlying coherence without reducing them to a collection of prescribed techniques. The families reveal recurring ways in which pedagogical support can be embedded within the learning environment while leaving considerable scope for creativity, responsiveness and adaptation. Individual activations frequently combine characteristics from several families. A card-based activation may incorporate movement, while a textbook-based activation may make use of artefacts, with spatial organisation supporting both. The boundaries between families therefore remain intentionally permeable, reflecting the dynamic nature of suggestopedic practice.

What remains constant is not the outward form of the activation but the pedagogical logic that underpins it. Whether support is distributed through language, materials or space, the purpose remains the same: to create conditions in which learners can participate successfully within a carefully designed learning environment. The families therefore illustrate not different methods of teaching but different ways of organising scaffolded participation across conscious and paraconscious planes of learning.

Viewed in this way, the concept of families provides more than a descriptive classification of classroom practice. It offers an analytical framework for understanding how contemporary Suggestopedia organises learning through intentionally designed environments. Rather than focusing on the activities themselves, the framework makes visible the pedagogical principles through which participation, communication, confidence and gradual independence are systematically supported.

## Design Principles of Type One Activations

The families described above illustrate how Type One activations are realised through different pedagogical resources. Despite this variation, a small number of recurring design principles can be identified. These principles are not intended as prescriptive rules but as analytical descriptions of the pedagogical conditions that enable supported participation within contemporary suggestopedic practice.

The first principle is the design of success. Type One activations are structured so that every learner can participate meaningfully from the outset. Rather than requiring independent production beyond current competence, support is continuously distributed throughout the learning environment through language, materials, interaction and teacher guidance. Participation therefore becomes possible before independence is expected, reflecting the pedagogical logic of scaffolding and the zone of proximal development.

The second principle is the cultivation of imagination. Type One activations invite learners into meaningful contexts where language and content are experienced through narrative, roles, movement, symbolic play and aesthetic framing. Imagination is therefore not an optional addition to learning but a condition that encourages exploration, emotional engagement and the construction of meaning.

A third principle is the anchoring of learning within a shared whole. Language and content are continually connected to a common text, dialogue or thematic world, allowing individual words, structures and experiences to retain their meaning within a broader context. This reflects Lozanov's principle of Global–Partial–Global, in which understanding develops through a continuous movement between the whole and its constituent parts (Lozanov 2009, pp. 59–60).

The fourth principle is the continuity of communication. Learners are invited to speak, listen, respond and interact from the very beginning of the learning process. Communication is supported rather than postponed, creating a shared social environment in which confidence, familiarity and linguistic competence develop together through participation.

Taken together, these design principles demonstrate that support is intentionally distributed across language, materials, space, interaction, and aesthetic experience, allowing participation to precede independent performance. Through this pedagogical design, learning unfolds simultaneously across conscious and paraconscious planes, enabling confidence,

communication, and understanding to develop as mutually reinforcing aspects of the same process.

## Discussion

The analysis presented in this paper suggests that Type One activations are best understood not as a collection of classroom activities but as a theoretical framework for understanding how contemporary Suggestopedia organises learning through intentionally designed environments. Viewing Type One activations in this way shifts attention from the visible forms of classroom practice to the pedagogical logic that underpins them. The central question is therefore no longer what activities are used, but how learning environments are designed to support participation, communication and the gradual development of independent performance.

One important implication concerns the concept of scaffolding. Within much educational research, scaffolding is primarily understood as support provided through interaction between teacher and learner. The present analysis suggests a broader perspective. In Type One activations, support is distributed throughout the learning environment itself. Language, materials, spatial organisation, movement, aesthetic framing and patterns of social interaction all function as mediating resources that make participation possible. Scaffolding therefore becomes not only an interactional process but also a characteristic of pedagogical design.

A second implication concerns the relationship between participation and learning. Rather than treating communication as a goal that follows the acquisition of knowledge, Type One activations position communication at the centre of the learning process itself. Learners are invited to participate from the outset within environments where support is continuously available. Participation therefore becomes one of the mechanisms through which language, confidence and understanding develop, rather than simply an outcome of prior learning.

The analysis also contributes to a broader understanding of Suggestopedia itself. The concept of families demonstrates that the same pedagogical principles can be realised through different forms of classroom design. Whether support is distributed through language, materials or space, the underlying pedagogical logic remains consistent. The families therefore function not as a classification of activities but as an analytical framework for understanding how contemporary suggestopedic practice organises supported participation while preserving flexibility, creativity and responsiveness to learners.

At the same time, the analysis highlights an important pedagogical consideration. When Type One activations are separated from the broader suggestive environment in which they are embedded, they risk being reduced to isolated techniques. Activities may retain their visible structure while losing the coherence that emerges from the interaction of aesthetic design, mediated participation, social communication and dual-plane learning. Their educational significance therefore lies not in the activities themselves but in the relationships they establish within the learning environment as a whole.

The present study has focused primarily on developing a theoretical account of Type One activations by bringing together perspectives from Suggestopedia and sociocultural theories of learning. As a conceptual paper, it does not examine empirical classroom data. The proposed framework should therefore be regarded as a basis for further investigation rather than as an empirically validated model. Future research may explore how these design principles are

realised across different suggestopedic traditions and examine their relationship to participation, confidence and language development in classroom practice.

Taken together, the analysis suggests that the effectiveness of Type One activations lies not in particular classroom techniques but in the intentional design of learning environments that enable supported participation across conscious and paraconscious planes of learning. In this sense, Type One activations provide more than a description of contemporary suggestopedic practice. They offer a theoretical language through which the pedagogical logic of contemporary suggestopedic practice can be analysed, discussed and further developed.

## Conclusion

This paper has sought to articulate the pedagogical logic of Type One activations and to position them within the theoretical framework of Suggestology and Suggestopedia. While widely recognised in Nordic suggestopedic practice, Type One activations have remained largely implicit in academic literature. By bringing together perspectives from Suggestopedia and sociocultural theories of learning, the paper has proposed a theoretical framework through which these activations can be understood as intentionally designed learning environments for supported participation.

The analysis has shown that Type One activations are characterised not by particular classroom activities, but by the way pedagogical support is organised. Through distributed scaffolding, shared contexts, aesthetic design and continuous communication, they create conditions in which learners are able to participate meaningfully from the outset while gradually enabling independent performance. In this way, learning develops through participation within a coherent learning environment rather than through the isolated performance of individual tasks.

The identification of recurring pedagogical families and underlying design principles has provided a language for describing how contemporary suggestopedic practice maintains coherence while remaining flexible and responsive to different learners and contexts. Rather than prescribing specific techniques, this framework makes it possible to analyse how language, materials, space, interaction and aesthetic experience work together to support learning across conscious and paraconscious planes.

More broadly, the paper suggests that the significance of Type One activations extends beyond Suggestopedia itself. By focusing attention on the intentional design of learning environments rather than on individual teaching techniques, the proposed framework contributes to wider discussions of scaffolding, participation, mediation and learning design. In this sense, Type One activations offer not only a way of understanding contemporary suggestopedic practice but also a conceptual perspective through which the relationship between support, participation and learning may be further explored.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that generative AI and AI-assisted writing technologies (including ChatGPT and Grammarly) were used to support language editing, proofreading and the improvement of clarity and readability. The author critically reviewed, revised and approved all AI-assisted suggestions and takes full responsibility for the final content of the manuscript.

The ideas, theoretical framework, pedagogical concepts, analyses, interpretations and conclusions presented in this paper are the author's own.

## References

Hammond, J., & Gibbons, P. (2005). Putting scaffolding to work: The contribution of scaffolding in articulating ESL education. *Prospect, 20*(1), 6–30.

Johnstone, C. (2026). *Trusting the learning mind*. Lärlusta.

Lozanov, G. (2005). *Suggestopaedia – Desuggestive teaching: Communicative method on the level of the hidden reserves of the human mind.* International Centre for Desuggestology.

Lozanov, G. (2009). *Suggestopedia–Reservopedia: Theory and practice of the liberating-stimulating pedagogy on the level of the hidden reserves of the human mind*. St. Kliment Ohridski University Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

**Contact email:** christina.johnstone13@gmail.com

# The DIAL Framework for AI-Supported Decision Making in Undergraduate Business Programs

Eduard Merc, University of Hawaiʻi at Mānoa, United States

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Artificial intelligence (AI) is rapidly transforming organizational decision-making processes, yet undergraduate business education has been slower to systematically integrate AI-supported decision making into its curricula. While students are increasingly exposed to analytics tools, there remains a gap in how these tools are structured within a coherent decision-making approach that combines data-driven insights with human judgment, ethical considerations, and contextual awareness. This paper proposes the DIAL (Decision Intelligence with AI Learning) Framework, which conceptualizes decision making as a multi-layered process consisting of input, process, and output components. At the input level, students develop data literacy and engage with AI tools to generate, interpret, and contextualize insights. The process level emphasizes the central role of human judgment and ethical awareness in critically evaluating AI-generated outputs, mitigating potential biases, and ensuring responsible and context-sensitive decision making. At the output level, the framework focuses on improving decision quality and enhancing business impact in real-world organizational settings. The framework is informed by teaching practices and industry experience in a diverse higher education environment in the State of Hawaiʻi, USA, highlighting the importance of culturally relevant, applied, and student-centered learning contexts. The DIAL Framework contributes to business education by offering a structured and scalable approach to integrating AI into decision-making instruction and provides a foundation for future empirical research on AI-supported learning in higher education and workforce preparation.

*Keywords:* artificial intelligence, decision intelligence, business education, AI-supported learning, data literacy

iafor
The International Academic Forum
www.iafor.org

## Introduction

Artificial intelligence is changing how organizations analyze information, generate alternatives, and make decisions. In business education, students now encounter generative AI alongside established analytics technologies such as spreadsheets, optimization tools, data visualization platforms, and business intelligence systems. The educational challenge is therefore no longer limited to whether students should use AI. A more consequential question is how students can learn to use AI while retaining the capacity to question data, evaluate outputs, recognize limitations, consider ethical consequences, and remain accountable for the final decision.

Research and professional guidance increasingly emphasize that effective educational use of generative AI requires more than technical access. Large language models can support learning and analysis, but their limitations create a need for critical thinking, fact checking, human oversight, and responsible use (Kasneci et al., 2023). UNESCO (2023) similarly advocates a human-centered approach to generative AI in education that addresses privacy, ethics, pedagogical appropriateness, and human agency. These concerns are particularly relevant in business programs, where students are preparing for roles in which AI-generated recommendations may influence financial, operational, customer, and strategic decisions.

This paper introduces the DIAL Framework as a conceptual and applied model for AI-supported decision making in undergraduate business education. DIAL is an acronym for Decision Intelligence with AI Learning and is operationalized through four connected elements: Data Literacy, Intelligent Augmentation, Applied Judgment, and Leveraged Outcomes. The framework is informed by the author's teaching practice in Hawai‘i and by classroom activities in which students compare their own analytical work with AI-supported outputs. The purpose is not to claim empirical validation, but to provide a practical structure that educators can use and future research can test.

## AI-Supported Decision Making in Business Education

Generative AI creates both opportunity and risk for management learning. It can accelerate information gathering, provide feedback, generate alternatives, and help learners explore business scenarios. At the same time, fluent AI output can appear authoritative even when it is incomplete, biased, or incorrect. Kasneci et al. (2023) argue that educational integration of large language models requires competencies that enable learners to understand both their capabilities and their limitations. Tlili et al. (2023) likewise identify educational opportunities while drawing attention to concerns involving truthfulness, privacy, manipulation, and responsible adoption.

For business education, the central issue is judgment. Management decisions rarely depend on a single technically correct answer. They involve context, uncertainty, competing objectives, stakeholder consequences, and ethical tradeoffs. Moser et al. (2022) caution against allowing algorithmic reckoning to substitute for human judgment and call for renewed attention to teaching judgment in management education. More recent work on management learning similarly emphasizes the potential for AI-supported experiential designs to strengthen reflection and judgment rather than simply automate student work (Johnson et al., 2026).

The DIAL Framework responds to this need by distinguishing AI tool proficiency from decision intelligence. Knowing how to prompt a model, build a dashboard, or run an analysis

is valuable, but those skills do not automatically establish whether an output should be trusted or how it should influence a decision. DIAL therefore treats AI as an augmentation mechanism inside a broader human decision process. Its educational objective is to make the movement from data to AI-supported insight, from insight to judgment, and from judgment to outcomes explicit and repeatable.

## The DIAL Framework

The DIAL Framework translates AI-supported decision making into four memorable components that can be embedded in business cases, analytics exercises, projects, and discussions. Although the accepted conceptual model can be viewed as input, process, and output layers, the four D-I-A-L components make those layers actionable for teaching.

Data Literacy is the foundation. Students first learn to read, interpret, and question the data available for a decision. They consider data quality, relevance, limitations, and the business question being asked before relying on an analytical or AI-generated result. This component is important because sophisticated AI cannot compensate for poor inputs, inappropriate measures, or misunderstood context.

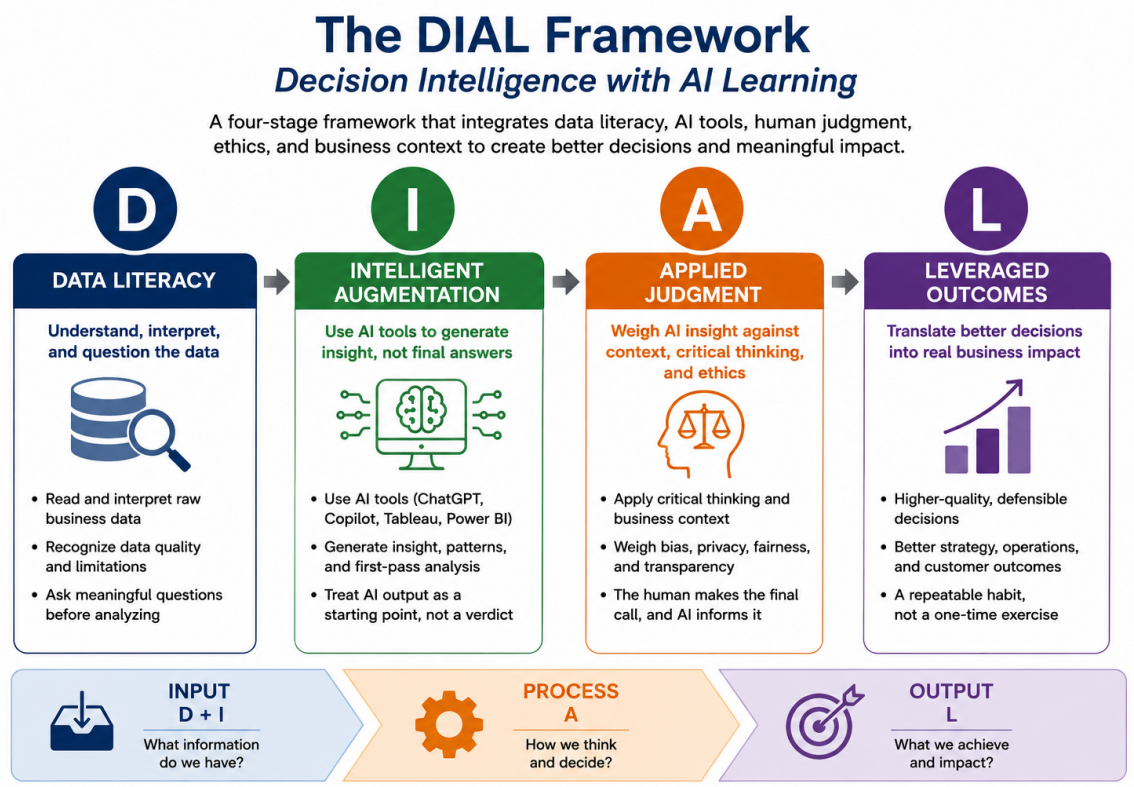
Intelligent Augmentation introduces AI as a source of additional insight rather than as the final decision maker. Students may use tools such as ChatGPT, Copilot, Tableau, Power BI, or other analytics technologies to identify patterns, summarize information, generate alternatives, or produce a first-pass interpretation. The central instructional principle is that AI output is a starting point, not a verdict. This framing aligns with human-centered approaches to educational AI that preserve human agency and critical evaluation (UNESCO, 2023).

Applied Judgment is the point at which students evaluate AI-supported insights against business context, critical thinking, and ethical considerations. Students are encouraged to ask why an AI output differs from their own analysis, what assumptions may explain the difference, what information may be missing, and what consequences could follow from acting on the recommendation. Bias, privacy, fairness, and transparency are considered within this judgment stage. Human oversight is therefore not an afterthought; it is the mechanism through which AI-supported information becomes a defensible decision.

Leveraged Outcomes completes the framework by connecting decision quality to business impact. Students move beyond identifying an interesting insight and explain what action should follow, why it is justified, and what organizational outcome it is intended to improve. Depending on the problem, this may involve strategy, operations, customer outcomes, resource allocation, or risk. The goal is to develop a repeatable decision habit in which students use AI without surrendering responsibility for the decision.

Together, the four components connect the framework's broader input-process-output logic. Data Literacy and Intelligent Augmentation support the input and insight-generation stage; Applied Judgment represents the critical human process; and Leveraged Outcomes focuses attention on decision quality and business impact.

**Figure 1**
*The DIAL Framework for AI-Supported Decision Making*




*Note.* Conceptual framework developed by the author.

## Classroom Applications of DIAL

The framework has been refined through classroom practice, although the examples presented here are instructional observations rather than results from a formal empirical study. Two applications illustrate how DIAL can be incorporated into business courses without requiring a separate AI course.

In an undergraduate Business Statistics course, students completed statistical scenarios manually using Excel, Solver, and Data Analysis tools. They then compared their results with AI-generated output addressing the same scenarios. Similarities and small discrepancies became the basis for discussion. Instead of asking only whether the AI was correct, students had to examine the data, identify possible reasons for differences, and determine which result they would rely upon. This activity reflects Data Literacy, Intelligent Augmentation, and Applied Judgment. The instructional value was located in the comparison process: disagreement between human and AI-supported analysis created an opportunity to question assumptions and defend a decision.

A second application occurred in an MBA Business Analytics course. Students created dashboards in Excel and Tableau from a shared dataset and compared their manually developed dashboards with an AI-supported dashboard generated from the same data. The comparison extended beyond numerical accuracy to the story communicated by the visualization: which measures were emphasized, what patterns were highlighted, and what the dashboard suggested decision makers should consider. This activity emphasized Applied Judgment and Leveraged Outcomes because students evaluated not only what the technology produced, but also how analytical choices shaped the interpretation and potential business action.

These examples demonstrate how DIAL can function as a shared vocabulary across different business subjects and levels. The framework does not prescribe a particular AI platform. Instead, it structures the questions students should ask before, during, and after using AI in a decision task.

## Limitations and Future Research

DIAL is currently a conceptual framework refined through the experience of a single instructor and has not yet been validated as an instructional model. The classroom examples are informal observations and were not designed as controlled studies; therefore, no causal claims about learning outcomes are made. This limitation is intentional and defines the next stage of the research agenda.

Future research should operationalize decision intelligence through measurable learning outcomes. Potential approaches include pre- and post-assessments, rubric-scored decision quality, student reflections, comparisons between AI-supported and non-AI-supported tasks, and replication across instructors, institutions, disciplines, and student populations. Research should also examine whether the framework helps students recognize inappropriate AI recommendations, articulate ethical concerns, and transfer AI-supported judgment to unfamiliar business problems. Hawaiʻi provides a valuable context for this work because its higher education environment connects local, international, and culturally diverse student populations; however, cross-institutional testing will be necessary before broader generalizability can be established.

## Conclusion

The rapid adoption of generative AI makes AI literacy necessary, but tool proficiency alone is insufficient preparation for business decision making. Students must also learn to question data, evaluate AI-generated insights, apply contextual and ethical judgment, and connect decisions to meaningful outcomes. The DIAL Framework provides a concise structure for teaching that progression.

DIAL organizes AI-supported decision making around Data Literacy, Intelligent Augmentation, Applied Judgment, and Leveraged Outcomes. Its central premise is that AI should augment rather than replace human decision making. The framework can be incorporated into existing statistics, analytics, information technology, and management activities by requiring students to compare, challenge, justify, and act upon AI-supported insights. At its present stage, DIAL should be understood as a conceptual and practice-informed framework rather than a validated instructional model. Its contribution is therefore both practical and prospective: it offers educators a usable structure today while establishing a foundation for empirical research on how AI-supported learning can strengthen decision quality, ethical awareness, and workforce preparation in business education.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that OpenAI's ChatGPT, a generative AI tool, was used to assist with organizing the manuscript, drafting portions of the text, and refining language for clarity and concision. The author developed the DIAL Framework, selected the classroom examples,

reviewed and revised the manuscript, verified the cited sources, and takes full responsibility for the accuracy, interpretation, and final content of the work.

## References

Johnson, M., Maitland, E., & Sofka, W. (2026). Developing judgement for business: An AI-based model of independent management learning. *Journal of Business Research, 204*, 115842. https://doi.org/10.1016/j.jbusres.2025.115842

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Moser, C., den Hond, F., & Lindebaum, D. (2022). Morality in the age of artificially intelligent algorithms. *Academy of Management Learning & Education, 21*(1), 139–155. https://doi.org/10.5465/amle.2020.0287

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*, 15. https://doi.org/10.1186/s40561-023-00237-x

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

# Development and Evaluation of a Gamified Digital Platform for Science Education in Secondary Schools

Nurseit Zhunissov, Khoja Akhmet Yasawi International Kazakh-Turkish University, Kazakhstan
Anuarbek Amanov, Khoja Akhmet Yasawi International Kazakh-Turkish University, Kazakhstan
Arypzhan Aben, Khoja Akhmet Yasawi International Kazakh-Turkish University, Kazakhstan

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

The integration of gamification into science education has emerged as an effective approach to enhance student engagement, motivation, and learning outcomes. This study investigates the application of gamification elements in teaching natural science subjects through the development of a specialized digital platform. The proposed platform incorporates key gamification components, including points, badges, levels, leaderboards, and instant feedback mechanisms, designed to foster interactive and student-centered learning. The research was conducted among 7th-grade students, where the developed platform was integrated into the instructional process. To evaluate its effectiveness, a quasi-experimental design was employed, including pre-test and post-test assessments, along with student motivation and engagement surveys. The evaluation framework was based on widely used educational measurement scales, such as Likert-scale questionnaires and performance-based assessment criteria, enabling both quantitative and qualitative analysis. The results demonstrated a significant improvement in students' academic performance and learning motivation compared to traditional teaching methods. Students who used the gamified platform showed higher test scores, increased participation, and greater interest in science subjects. Additionally, the platform contributed to the development of critical thinking and problem-solving skills by encouraging active involvement in the learning process. Furthermore, the findings align with existing research highlighting the positive impact of gamification in education, particularly in STEM disciplines. The use of structured evaluation scales allowed for a comprehensive assessment of both cognitive and affective learning outcomes. In conclusion, the developed gamification-based platform proved to be an effective tool for enhancing the quality of science education.

*Keywords:* gamification, natural science education, digital platform, student motivation, STEM education

iafor
The International Academic Forum
www.iafor.org

## Introduction

One of the main tasks of education systems in the 21st century is to develop students' functional literacy and research skills through the effective use of information and communication technologies. The digitalization of education in the Republic of Kazakhstan has become a priority direction of state policy, reflecting a broad international trend towards the use of technologies in the learning environment. In accordance with this policy direction, schools in the country are increasingly introducing electronic resources, virtual laboratories, interactive platforms and gaming technologies into everyday teaching. Natural sciences—physics, chemistry and computer science—are the main focus of this modernization work, as they require students to acquire not only theoretical knowledge, but also to develop practical skills, scientific thinking and a logical-analytical approach. These competencies are important for students who intend to study or build a career in science, technology, engineering and mathematics (STEM) and are increasingly being emphasized in national curricula. However, a significant part of educational research shows that students often perceive these subjects as complex, abstract and relatively uninteresting, especially when teaching is based on lectures and memorization. Common barriers include limited access to hands-on lab experiences, lack of immediate feedback, and lack of individualized learning pace that takes into account students' varying levels of knowledge. These circumstances create a need to redesign the learning process in a more engaging, interactive, and learner-centered way. One approach that has gained significant popularity in education in recent years is the gamification approach. The gamification approach involves applying game elements and game mechanics—such as points, badges, levels, and structured feedback—to non-game contexts, including formal education, with the goal of increasing user engagement and motivation. Unlike full-scale educational games or game-based learning environments, the gamification approach does not require the creation of a completely fictional game world; instead, it overlays familiar game mechanics onto existing course content and tasks. In an educational environment, the gamification approach aims to increase student motivation by incorporating game elements into the learning material: as students progress through the course, they advance to the next level, earn rewards, and reinforce their learning by completing structured tasks. This approach aims to make the learning process more enjoyable and productive by transforming task completion into a process with visible, incremental progress and indicators of achievement.

Gamified digital platforms for science education offer a number of tangible benefits beyond increasing interest in the subject. They allow students to virtually complete lab work when access to a physical lab is limited, take quizzes in a gamified format that reduces test anxiety, and track their progress using visible progress indicators. In addition, such platforms allow teachers to automatically assess student performance, analyze results, and identify students who may need additional support, thereby reducing the administrative burden associated with manual review and progress monitoring.

This article reviews the theoretical foundations of the development of a game-based digital platform for secondary schools, presents the structure and functional architecture of the platform, and describes a system for evaluating its pedagogical effectiveness.

## Literature Review

Currently, the rapid development of digital technologies in education has greatly influenced the widespread use of new pedagogical approaches based on the use of game elements in the

teaching process. Gamification in teaching natural sciences is considered an effective tool for increasing students' motivation to learn, cognitive activity and subject competence.

Indiaka et al. (2025), who studied the effectiveness of digital gamification in teaching biology in general education institutions, found that game-based digital teaching methods significantly improve students' level of mastery of biological concepts and motivation to learn. According to the author, digital games increase students' interest in the learning material, contribute to their active participation in the lesson and long-term memory of knowledge.

Moral-Sánchez et al. (2022), who studied the teaching experience of mathematics and STEM, considered that combining STEM and gamification methods in teaching geometry allows students to develop problem-solving, critical thinking and collaborative skills. This is believed to prove the importance of gaming technologies in not only increasing interest, but also in developing 21st century skills.

In their work, Ateş and Kölemen (2025), who examined the issue of teaching science through mobile devices, proved that mobile educational platforms incorporating gaming elements showed high results in developing students' self-regulated learning abilities. The authors noted that digital platforms contribute to students' self-planning of the learning process, monitoring their academic achievements, and increasing their responsibility.

In a study conducted in the field of science teacher training, Franco-Mariscal, Cebrián-Robles, and Rodríguez-Losada (2023) examined the effectiveness of using electronic rubrics to build teachers' competence in evaluating gamification resources. The results of the study showed that teachers' high-quality development and evaluation of digital gaming resources increases the effectiveness of the educational process.

Katanosaka et al. (2023), who developed an online game-based platform for teaching physics, analyzed the platform's architecture, functionality, and impact on students' academic achievement. The researchers found that virtual experiences, graded tasks, and reward systems facilitated students' understanding of physical phenomena and improved learning outcomes.

Alarcon et al. (2025), who evaluated the effectiveness of the ABCCI serious game program used in science teacher training, found that the usability and gamification elements of the game positively influenced the development of professional competence. This study, published in the IEEE Access journal, demonstrates that digital games play an important role not only for students but also in the professional training of teachers.

Carvalho et al. (2024) developed a gamified digital platform aimed at organizing extracurricular activities for elementary and secondary schools. Preliminary results showed that game mechanics increase students' active participation in school life and enhance interaction in the digital environment. This study demonstrates that digital platforms are not only effective in academic learning, but also in educational and social work.

In a systematic literature review on the use of board games in science education, Hashim et al. (2023) analyzed the impact of gaming technologies on the quality of education. They showed that gaming methods facilitate the acquisition of subject knowledge and contribute to the development of students' creativity, communication skills, and a culture of collaboration. The study emphasizes that the combined use of digital and traditional games yields effective results.

Khattib and Alt (2024) evaluated the effectiveness of digital gamification in science education using the CoSpaces Edu platform based on a quasi-experimental study. The results of the study show that the academic achievement, spatial thinking skills, and digital competences of students in the experimental group were significantly higher than those in the control group. This demonstrates the potential of digital platforms to improve the quality of education.

In their systematic literature review on the use of gamification in secondary education, Vrcelj et al. (2023) analyzed numerous studies published in recent years and concluded that gamification consistently improves learning motivation, engagement, and learning outcomes. They argued that pedagogical design, user interface, and the coherence of learning objectives should be among the key requirements in the development of digital platforms.

The analysis of the reviewed scientific works proves that game-based digital platforms are highly effective in teaching natural sciences. Many studies have proven that gamification, in addition to increasing students' interest in the subject, contributes to the organization of their own learning, the development of research skills, and the formation of digital competence. At the same time, it is noticeable in the literature that there is still a lack of research on complex solutions that combine several areas of natural sciences (chemistry, physics, computer science) within one digital platform. In this regard, the development of a complex game-based digital platform for secondary schools and the assessment of its effectiveness are relevant scientific issues for the modern education system.

## Materials and Methods

### Platform Overview

The main material for the study is the digital educational game platform EduGame.kz, developed for general secondary schools in the Republic of Kazakhstan. The platform is a web-based system designed to increase student motivation, monitor academic progress, and support the development of subject-specific competencies. Its interface is designed in accordance with the basic principles of game design and provides interactive learning in the natural sciences: physics, chemistry, and computer science.

The platform is not intended to replace traditional classroom teaching, but rather to expand and enhance it as an additional learning environment. Teachers have full control over the materials provided to students, the pace of task completion, and the activation or deactivation of individual game elements such as leaderboards, which allows for flexible adaptation to the needs of different grade levels and age groups.

### Gamification Module

The most important of the main components of the digital platform developed during the study is its gamification module. This module helps to increase students' motivation to learn and ensure their active participation in the learning process. It is designed to improve academic achievement in science subjects. By integrating gamification mechanics with educational content, it not only increases students' cognitive interest, but also maintains their activity over time. In developing the module, the research team uses the basic framework proposed by Deterding, Dixon, Khaled, and Naik (2011), which defines modern game mechanics as, in particular, game design elements and the use of design elements in contexts; the work of Kapp (2012) on game-based methods and strategies for teaching and learning; and the taxonomy of

dynamics, mechanics, and components developed in the works of Werbach and Hunter (2015). These studies influenced decisions on how to prioritize, organize, and balance them in accordance with the developmental characteristics of secondary school students and the requirements of the science curriculum. Gamification elements in the platform are automatically activated when students complete learning tasks and the system is designed to provide immediate feedback on student progress, which, according to the literature reviewed above (e.g., Ateş & Kölemen, 2025; Katanosaka et al., 2023), is crucial for maintaining interest and supporting self-regulated learning.

## Key Elements and Functions of the Gamification Module

The game module developed for the study includes 10 key elements, briefly described in Table 1 below, each of which performs a specific pedagogical function in the overall system.

**Table 1**
*Core Elements of the EduGame.kz Gamification Module*

| Element | Function and Capabilities |
|---|---|
| **Points** | Awarded automatically for every correctly completed task. Points function as the primary quantitative indicator of learning activity, allowing both students and teachers to monitor engagement and achievement over time. |
| **Badges** | Digital badges are granted upon reaching specific learning milestones (e.g., completing a topic, achieving a streak of correct answers). Badges serve as visible symbols of competence and are intended to strengthen students' intrinsic motivation. |
| **Levels** | Students advance through sequential levels as they accumulate experience points. Each new level unlocks tasks of increasing complexity, ensuring that challenge scales with demonstrated competence. |
| **Leaderboard** | Displays student rankings based on performance, fostering a spirit of friendly competition. Teachers may enable or disable this feature depending on classroom needs, to avoid demotivating lower-performing students. |
| **Quests** | Thematic learning missions aligned with specific topics in physics, chemistry, and computer science. Quests require the sequential completion of several related tasks, encouraging sustained engagement with a topic. |
| **Daily Challenges** | Short, time-limited tasks designed to encourage regular, spaced learning activity and to build consistent study habits. |
| **Progress Bar** | A visual indicator that displays a student's cumulative progress through a course or module, helping learners track how much material remains. |
| **Immediate Feedback** | Following each task, students receive instant information about the correctness of their answer along with explanatory comments, supporting formative assessment and timely correction of misconceptions. |
| **Rewards** | Additional privileges, certificates, or virtual prizes granted upon reaching defined performance thresholds, reinforcing sustained achievement. |
| **Adaptive Tasks** | The difficulty level of subsequent tasks is automatically adjusted based on a student's prior performance, supporting individualized learning trajectories. |

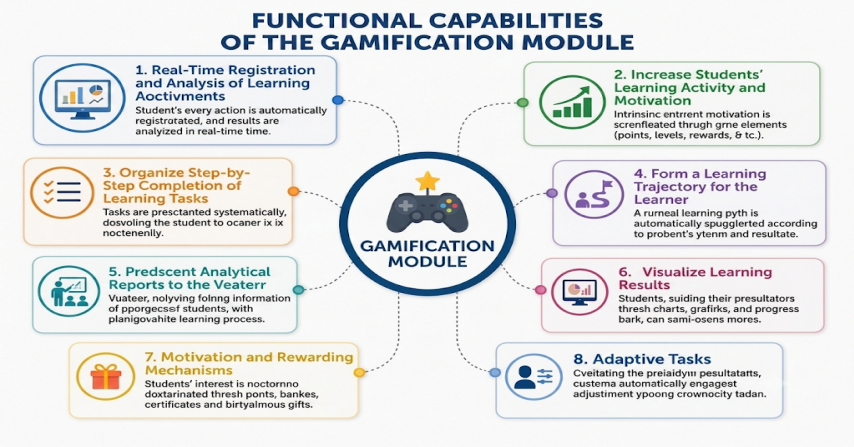
## Module Functionalities

This module was subsequently piloted to assess the pedagogical effectiveness of the platform during the practical phase of the study. The developed gamification module provides the following functionality, schematically illustrated in Figure 1: automatic tracking of student

progress across topics and tasks; calculation and real-time display of accumulated points, badges, and level status; teacher dashboards that summarize class and individual performance; and settings that allow teachers to customize the visibility of competitive elements, such as leaderboards, to suit the needs of a particular class.

**Figure 1**

*Functional Capabilities of the EduGame.kz Gamification Module*




## Research Methodology

A quasi-experimental research design combining pre- and post-assessment with student motivation and engagement measurements was used to evaluate the effectiveness of the proposed platform. The study was conducted among 7th grade students, for whom the platform was introduced into the regular curriculum of the target science subjects. This age group was chosen because 7th grade represents the stage when students begin to systematically study individual physics, chemistry, and computer science courses. It is a suitable context to study the impact of the platform on basic knowledge acquisition and interest in specific subjects.

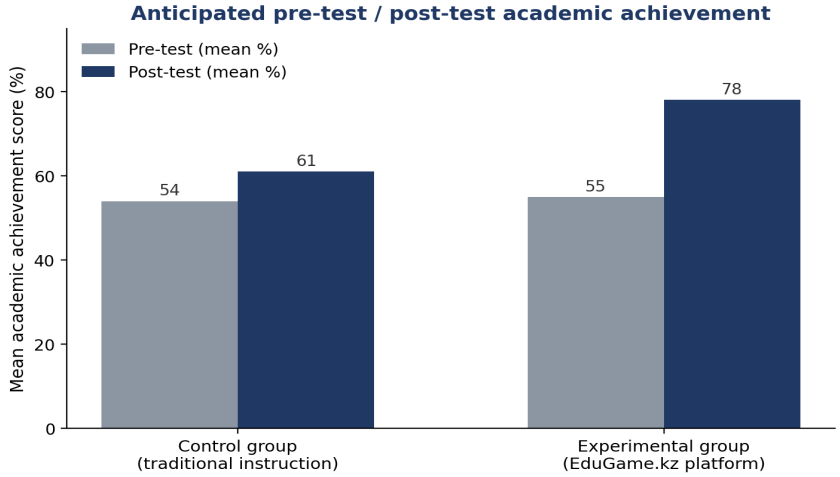
The assessment tools used in this study are based on widely accepted measurement approaches in education, including Likert-scale questionnaires for assessing motivation and attitudes and indicator-based assessment criteria for assessing academic performance. This combination of tools was chosen to quantitatively analyze learning outcomes and qualitatively understand students' subjective experiences in a game-based learning environment.

## Analysis and Expected Results

The proposed game-based digital platform will be evaluated through a pilot study aimed at comprehensively assessing the impact of the proposed game-based digital platform on the acquisition of natural science subjects (Physics, Chemistry, and Computer Science), motivation to learn, and overall academic performance of secondary school students.

In the next stage of the study, experimental and control groups will be created, and students' knowledge will be assessed through pre- and post-testing. This design will allow us to identify changes in academic performance as a result of using the platform. At the same time, the results of the experimental and control groups will be compared, and the statistical significance of any differences observed will be assessed using appropriate inferential methods. Figure 2 shows an illustrative projection of the expected results of this comparison based on the trends described in the game-based literature discussed.

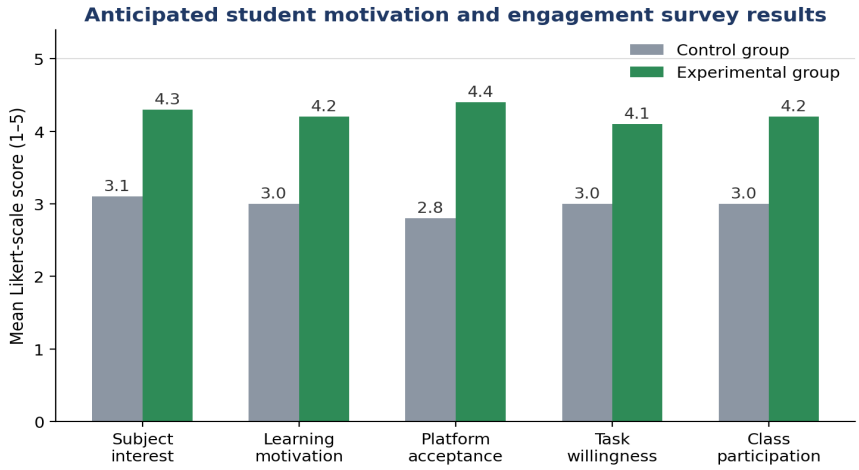
**Figure 2**
*Anticipated Pre-test / Post-test Academic Achievement*




*Note*. illustrative projection, not empirical data

A Likert-scale survey is planned to assess students' motivation to learn. The survey results will be used to analyze students' interest in the subject, their acceptance of the digital platform, their readiness to complete assigned tasks, and their overall level of engagement in the learning process. The survey questions are designed to capture cognitive aspects of motivation, such as perceived competence and value, as well as emotional aspects, such as enjoyment and reduced anxiety related to the subject. The expected survey results for these aspects are shown in Figure 3.

**Figure 3**
*Anticipated Student Motivation and Engagement Survey Results*

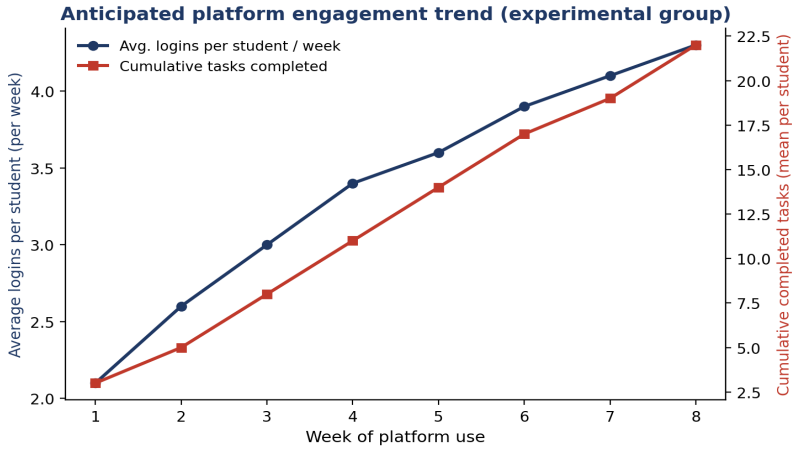



*Note*. illustrative projection, not empirical data

The platform's built-in analytics module automatically tracks users' learning activity. This analysis looks at metrics such as login frequency, number of tasks completed, points earned, badges earned, tasks completed, and overall learning progress. This data allows us to assess students' engagement on the digital platform and explore the relationship between engagement and academic performance. Figure 4 shows the expected engagement trend over the 8-week trial period, with both login frequency and total number of tasks completed gradually increasing as students become more familiar with the platform.

**Figure 4**

*Anticipated Platform Engagement Trend Over an Eight-Week Trial*




*Note*. illustrative projection, not empirical data

The platform will conduct a comparative analysis of the impact of gamification elements such as point systems, levels, badges, leaderboards, tasks, and feedback on student motivation and academic performance. In addition, the relationship between the frequency of use of each element and student academic performance will be examined, allowing the research team to determine which mechanisms contribute most to the observed improvements and which may require further improvement. Descriptive statistical methods, including means, standard deviations, and percentage distributions, are planned to process the collected experimental data. Depending on the characteristics of the data, student t-tests or non-parametric statistical methods will be used to identify differences between groups. In addition, correlation analysis is proposed to assess the relationship between learning activity and academic performance, which will allow us to understand whether high-level interaction with the platform is systematically associated with improved academic performance. Tables, histograms, line charts, and graphs describing academic performance have been developed to visually present the research results. These visualizations are intended to support a comprehensive assessment of the platform's effectiveness and to facilitate the interpretation of the results, both for the research team and for teachers and administrators who may subsequently consider implementing the platform.

The expected outcomes of using a gamified digital platform are to increase students' interest in science, increase their motivation to learn, increase their engagement in learning tasks, and contribute to improved academic performance. The results of the planned experiment provide scientific justification for the pedagogical effectiveness of the proposed platform. It is expected that they will serve as a basis for developing practical recommendations for its integration into the school curriculum. A comparative review of traditional teaching approaches with gamification-based learning, shown in Figure 5, highlights several key differences with respect to the expected outcomes of this study. Traditional science teaching is typically based on passive transmission of knowledge, providing delayed feedback through periodic testing, and relatively limited opportunities for individualized learning. In contrast, gamification-based learning implemented through the EduGame.kz platform is characterized by active, task-oriented participation, immediate feedback after each learning activity, visible and gradual progress indicators, and an adaptive task sequence that takes into account the individual performance of students.

**Figure 5**

*Comparison of Traditional Instruction and Gamification-Based Instruction*

COMPARISON OF TRADITIONAL LEARNING AND GAMIFICATION-BASED LEARNING

| Aspect | Traditional Learning | Gamification-Based Learning |
|---|---|---|
| Student's role | Passive listener | Active participant, researcher |
| Motivation | Extrinsic (grades, punishment) | Intrinsic and extrinsic (interest, excitement, satisfaction, reward) |
| Feedback | Insufficient | Immediate feedback during the task |
| Type of tasks | Traditional assignments, tests | Quests, missions, game scenarios |
| Assessment type | Summative assessment | Continuous assessment (process + result) |
| Engagement level | Average | High |
| Knowledge retention level | Low | High, long-term memory |
| Skills development | Limited | Critical thinking, problem-solving, creativity, teamwork |

## Conclusion

The study presents the structure and functionality of a game-based digital platform for teaching natural sciences in secondary schools. The platform integrates gamification elements aimed at activating the learning process, such as points, levels, badges, tasks, leaderboards, progress indicators, and instant feedback, and describes the methodological framework for integrating these elements with educational content.

An evaluation model was developed to evaluate the proposed platform, which includes indicators of learning achievement, learning motivation, learning activity, and platform usage. In addition, the data collection and statistical analysis methods used in the planned pilot study were described. Together, these methods are designed to provide a comprehensive assessment of the pedagogical effectiveness of the platform.

In the next stage of this study, the proposed digital platform will be tested in real school practice, the results of the experimental and control groups will be compared, and the impact of gamification elements on learning outcomes will be determined through statistical analysis. The expected results are expected to scientifically substantiate the conclusion that the use of game-based digital technologies increases students' interest in natural sciences, increases motivation to learn, and improves the overall quality of education. In general, the proposed game-based digital platform offers a promising pedagogical solution for improving the effectiveness of natural science teaching and can serve as a basis for further research and practical studies aimed at integrating it into the educational process of general secondary education institutions.

## Acknowledgements

This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant № AP 26193791). The authors would like to thank the Committee for the financial support that made this study possible. The authors also express their sincere gratitude to all colleagues and reviewers who contributed to improving the quality of this study.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

Generative artificial intelligence technology was used only to assist in the translation and linguistic processing of this manuscript. In particular, the ChatGPT tool (OpenAI) was used to improve the English translation, grammar, and readability of the text. This artificial intelligence tool was not used to generate research ideas, methodology, data analysis, results, or conclusions.

## References

Alarcon, D. A. U., Talavera-Mendoza, F., & Paucar, F. H. R. (2025). Evaluation of the serious game ABCCI: usability and gamification in teacher training for science education and environmental sustainability. *IEEE Access*, *13*, 90601–90611.

Ateş, H., & Kölemen, C. Ş. (2025). Gamifying mobile-based science education: enhancing self-regulated learning skills in middle school students. *Education and Information Technologies*, *30*(17), 24137–24171. https://doi.org/10.1007/s10639-025-13731-1

Carvalho, A. R., Ferreira, L., Gonçalves, L., Santos, C., Almeida, J., Ribeiro, N., ... & Pedro, L. (2024). Developing a gamified digital platform to promote extracurricular activities in basic and secondary schools–preliminary findings. *Educational Media International*, *61*(1-2), 103–116.

Cebrián-Robles, D., Cano-Iglesias, M. J., España-Naveira, P., & Franco-Mariscal, A. J. (2023). Integrating Corubric with Gamification of Resources in Science Education.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011, September). From game design elements to gamefulness: defining "gamification". In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9–15). https://doi.org/10.1145/2181037.2181040

Hashim, N. H., Harun, N. O., Ariffin, N. A., & Abdullah, N. A. C. (2023). Gamification using board game approach in science education-a systematic review. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *33*(3), 73–85.

Indiaka, S., Mwangi, W., & Sanyanda, J. (2025). Effectiveness of Digital Gamification in Biology Concept Comprehension and Learner Motivation in Secondary Schools. *Journal of Learning for Development*, *12*(3), 519–530. https://doi.org/10.56059/jl4d.v12i3.1595

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.

Katanosaka, T., Khan, M. F. F., & Sakamura, K. (2023, March). A physics learning system using gamification for high-school students. In *2023 11th International Conference on Information and Education Technology (ICIET)* (pp. 167–171). IEEE.

Khattib, H., & Alt, D. (2024). A quasi-experimental study on the advantages of digital gamification using CoSpaces Edu application in science education. *Education and Information Technologies*, *29*(15), 19963–19986.

Moral-Sánchez, S. N., Sánchez-Compaña, M. ª. T., & Romero, I. (2022). Geometry with a STEM and gamification approach: A didactic experience in secondary education. *Mathematics*, *10*(18), 3252. https://doi.org/10.3390/math10183252

Vrcelj, A., Hoic-Božic, N., & Dlab, M. H. (2023). Use of Gamification in Primary and Secondary Education: A Systematic Literature Review. *International journal of educational methodology*, *9*(1), 13–27.

Werbach, K., & Hunter, D. (2015). *The gamification toolkit: dynamics, mechanics, and components for the win*. University of Pennsylvania Press.

**Contact email:** nurseit.zhunissov@ayu.edu.kz

# A Comparative Study on Oral Communication Strategy Use Among Vietnamese Non-English Majors

Huyen Nhi Vo, Ton Duc Thang University, Vietnam
Huu Tai Nguyen, Ton Duc Thang University, Vietnam

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This paper investigated differences in the use of Communication Strategies among Vietnamese non-English majors of different English proficiency levels at the Creative Language Center, Ton Duc Thang University. A total of 101 undergraduates, at elementary and intermediate levels, answered a 58-item questionnaire, an adapted version of the Oral Communication Strategies Inventory (OCSI) suggested by Nakatani (2006), divided into two sections of Speaking and Listening. In terms of Speaking, "Message reduction and alteration" was the most frequently used category at both levels, with no significant difference between them. Significant differences were found in only two speaking categories: "Social affective", used more frequently by the intermediate group, and "Message abandonment", used more frequently by the elementary group. Regarding Listening, the highest-ranked category was "Word-oriented" for the intermediate group and "Negotiation for meaning while listening" for the elementary, though elementary means were higher in both. The differences between two levels were recorded in "Fluency-maintaining", "Less-active listener" and "Negotiation for meaning while listening", as the higher level more frequently used "Fluency-maintaining" and the lower-level group resorted to the other two strategies more often. Proficiency was therefore associated with differences in three of the seven listening categories but only two of the eight speaking categories, and in five of the seven listening categories it was the elementary group that returned the higher mean. The findings offer teachers and learners a basis for selecting strategies appropriate to each level, in order to compensate for gaps in language ability and to communicate more effectively.

*Keywords:* communication strategies, English proficiency, non-English majors, oral communication

iafor
The International Academic Forum
www.iafor.org

## Introduction

Oral communication in a foreign language seems to be as important as, if not more important than, other skills, such as reading and writing (Oweis, 2013). In Vietnam as in many other countries, English is an official subject in the national education system, and the ability to communicate in English is required by many employers. However, a majority of Vietnamese tertiary students struggle with their oral communication skills in English despite significant efforts to revolutionise the Vietnam education system (Phuong & Tran, 2021; Tran, 2013).

It is widely acknowledged that when learners suffer from communication breakdowns (Rabab'ah, 2003), they manage to overcome them by employing what are known as Communication Strategies (CSs) in order to fill in the gap between their communication intentions and the linguistic abilities they have. Rabab'ah (2016) concluded that it is necessary for second language learners to use interactional CSs in second language communication in order to first achieve their communicative goals, negotiate meaning, and finally improve their communicative ability. Therefore, CSs are believed to play an integral part for students to cope with speaking and listening problems in the process of language acquisition.

Some research has been conducted in Vietnam on the use of CSs among tertiary students. Among English majors, T. T. Q. Bui and Intaraprasert (2012, 2013) identified influences on strategy use and Le et al. (2020) found no differences across academic levels; among non-English majors, H. C. Nguyen (2023) and T. T. H. Nguyen et al. (2022) reported benefits of strategy use for communicative competence. These studies are reviewed in detail below.

Despite numerous research conducted in Vietnamese tertiary education, there has been a lack of findings which could compare the differences in the use of CSs among non-English majors of different levels of English proficiency. As a majority of Vietnamese students are non-English majors who have different English proficiency levels, the current paper, therefore, focuses on the comparison of how students of different levels of proficiency, namely elementary-level and intermediate-level, use common CSs, particularly looking at the frequency and the correlation between the CSs used and the corresponding language proficiency.

The research questions were, therefore, formulated as:

1. How do elementary and intermediate non-English majors use CSs in Vietnamese tertiary education?
2. What are the correlations between English language proficiency and the CSs use?

## Literature Review

### Definitions of Communication Strategies

CSs are defined as strategies a person having difficulties in achieving a communication goal uses intentionally to resolve those difficulties (Corder, 1983; Faerch & Kasper, 1980; Nugroho, 2019); learners held back by an incomplete linguistic system, or by limitations in knowledge and the overload of working memory, employ them to bridge the gap (Mitchell & Myles, 1998; Oxford, 1990). Oweis (2013) locates CSs in the individual language user, who experiences the problem and decides on a strategic plan for its solution.

A second perspective treats CSs as interactional rather than individual. They are seen as the mutual effort both speakers make to agree on meaning (Tarone, 1980, 1981) and as playing a key role in the effectiveness of communication (Bialystok, 1990; Canale, 1983). Le et al. (2020) combine the two views, defining CSs as verbal or non-verbal means employed by two or more interlocutors to negotiate meaning or resolve difficulties in both speaking and listening, so as to reach agreement on a communication goal.

## Classification of Communication Strategies

In respect of the classification of CSs, it also appears that different classification systems have been used. There is no complete agreement on what CSs are, how many exist, or how they should be classified, since past research has varied in setting, population, method and focus. The classifications of oral CSs (T. T. H. Nguyen et al., 2022) can be found in Table 1 below.

**Table 1**
*Classifications of Oral Communication Strategies*

| | |
|---|---|
| Faerch and Kasper (1980) | 1. Reduction strategies: Formal reduction; Functional reduction<br>2. Achievement strategies: Compensatory strategies or non-cooperative strategies; Retrieval strategies |
| Tarone (1981) | 1. Paraphrase: Approximation, word coinage, circumlocution<br>2. Borrowing: Literal translation, language switch, appeal for assistance, mime.<br>3. Avoidance: Topic avoidance, message abandonment. |
| Bialystok (1990) | 1. First language-based strategies<br>2. Second language-based strategies<br>3. Non-language strategies<br>4. Analysis-based strategies<br>5. Control-based strategies |
| Dörnyei (1995) | 1. Avoidance or reduction strategies<br>2. Message abandonment<br>3. Topic avoidance<br>4. Achievement or compensatory strategies<br>5. Approximation<br>6. Use of all-purpose words<br>7. Word coinage<br>8. Use of non-linguistic means<br>9. Literal translation<br>10. Foreignizing<br>11. Code switching<br>12. Appeal for help<br>13. Time-gaining strategies<br>14. Use fillers or hesitation devices |

| | |
|---|---|
| Corder (1983) | 1. Message adjustment strategies<br>2. Resource expansion strategies |
| Rabab'ah (2003) | 1. First language-based strategies (literal translation and language switch)<br>2. Second language-based strategies (avoidance strategies, word coinage, circumlocution, self-correction, approximation, mumbling) |
| Nakatani (2006) | Speaking<br>1. Social affective<br>2. Fluency-oriented<br>3. Negotiation for meaning while speaking<br>4. Accuracy oriented<br>5. Message reduction and alteration<br>6. Nonverbal strategies while speaking<br>7. Message abandonment<br>8. Attempt to think in English<br>Listening<br>1. Negotiation for meaning while listening<br>2. Fluency- maintaining<br>3. Scanning<br>4. Getting the gist<br>5. Nonverbal strategies while listening<br>6. Less active listener<br>7. Word-oriented |

*Note.* Adapted from T. T. H. Nguyen et al., 2022

**Communication Strategies Used and Language Proficiency**

Several papers have explored how English learners use CSs in various contexts. Mirzaei and Heidari (2012) explored the use of oral CSs by fluent and non-fluent L2 speakers at Shahrekord University, Iran. They noted that because fluency is related to communicative effectiveness, it has received considerable attention in L2 research. The ability of less proficient L2 learners to maintain the conversation flow is a worthwhile capacity because they have the chance to receive more language input by keeping a conversation going and thus to improve their L2 proficiency.

Similarly, Jindathai (2016) observed that students at the College of General Education and Languages, Thai-Nichi Institute of Technology, Bangkok, Thailand, regardless of their proficiency level, often used "Message reduction and alteration strategies" to address speaking challenges, and frequently negotiated for meaning while listening to overcome oral difficulties. These students also relied on "Social affective, Circumlocution, and Fluency-oriented strategies", to handle communication difficulties. More research was done to highlight the correlation between CSs use and language proficiency. La-Ortip (2003) reported that less fluent English majors at Rajabhat Institute Nakhon Ratchasima used CSs more frequently overall. More fluent students used English language-based strategies and avoided none, while weaker speakers used non-linguistic, L1-based and avoidance strategies more often. Chen (2009) and

Mirzaei and Heidari (2012) found that fluent speakers commonly used "Social affective, Fluency-oriented, Negotiation for meaning, and Accuracy-oriented strategies", while less fluent speakers favoured "Message reduction and alteration, and abandonment strategies". Ansarin (2003) studied 50 first-year and 50 final-year students at the University of Tabriz, Iran, and found no significant overall difference in CS use among Iranian learners of varying proficiency levels, although lower-level learners gave up more easily, while intermediate learners employed more approximation and general words to maintain interaction. Ting and Lau (2008), studying 14 pairs of Malaysian ESL learners, found that proficient speakers used lexical repetition and other discourse-based strategies to maintain conversations, while less proficient learners relied far more heavily on improvisation and avoidance, repeatedly restructuring their messages and abandoning them when restructuring failed; clarification requests and comprehension checks were used at a similar rate by both groups.

**Communication Strategies Studied in Vietnam**

Research in Vietnam has focused mainly on the relationship between learners' language proficiency and their use of CSs. T. T. Nguyen and Nguyen (2016) examined intermediate non-English majors through recordings of a group discussion task and informal interviews, and found that "Foreignizing", "Approximation", "Circumlocution", "Repetition" and "Expressing non-understanding" were not used at all, the most common strategies being fillers and hesitation devices, followed by self-repair, self-repetition and code-switching. T. T. H. Nguyen et al. (2022) found message reduction and alteration and accuracy-oriented techniques to be the least frequently used among tertiary students, and H. C. Nguyen (2023) reported that speaking strategies had a beneficial influence on the communicative competence of pre-intermediate undergraduates. Tuan and Mai (2015) point to limited practice in spoken English, and Trinh and Mai (2018) to the difficulty of running pair and group work with learners unfamiliar with such pedagogies.

Among English majors, Le et al. (2020) found fluency-oriented, message reduction and alteration and negotiation for meaning while speaking to be the most commonly employed speaking strategies, and "Negotiation for meaning while listening", "Non-verbal while listening" and "Scanning" the most common listening strategies; no significant differences emerged across academic levels, which the authors read as evidence of the universality of these strategies. T. T. Q. Bui and Intaraprasert (2013) found a strong association between attitudes towards speaking English and CS use, but only a minor one between CS use and exposure to oral communication.

Examining a range of learner factors, M. H. Bui (2016) found significant relationships between CS use and motivational orientation, English language proficiency, academic major and university setting, with qualitative data further linking Coping-CS and Developing-CS use to gender, cultural background and prior experience of language teaching and learning. In a large-scale study, T. T. Nguyen and Nguyen (2016) also surveyed teachers' views on the treatment of CSs in ESL textbooks and on whether CSs should be incorporated into the teaching of spoken English.

# Materials and Methods

## Research Setting

The study took place at the Creative Language Center, Ton Duc Thang University, which provides English programmes for non-English majors at levels ranging from elementary to advanced. Placement tests are used to assign students to appropriate classes and so minimise differences in language ability within a class. Each level covers all macro skills, with additional practice materials depending on the target exam, such as the Preliminary English Test (PET) or the International English Language Testing System (IELTS).

## Participants

The participants, mostly second-year students, were drawn from two groups of proficiency, elementary and intermediate, and completed the same questionnaire. The numbers of students in each group were 50 and 51, respectively. Participants were selected in two stages: students were first assigned to a level by the Center's placement test, and their class teachers were then consulted to identify those whose classroom performance clearly matched the CEFR descriptors for that level. The sampling was therefore purposive rather than random.

## Instrument and Data Analysis

The survey used in the present study was adapted from the Oral Communication Strategies Inventory (OCSI) suggested by Nakatani (2006) including 58 items, divided into 8 categories for coping with speaking problems and 7 for coping with listening problems, all of which are named in Tables 2 and 3. The items are translated into Vietnamese for students' better understanding and convenience as answering the questions.

In terms of data analysis, descriptive statistics Mean (M) of each item and category were processed. The following 5-point Likert scale was used to interpret the data: Never ($M \leq 1.80$); Rarely ($M$ = 1.81–2.60); Sometimes ($M$ = 2.61–3.40); Often ($M$ = 3.41–4.20); Always ($M$ = 4.21–5.00). Differences between the two proficiency levels, elementary (E) and intermediate (I), were examined using Pearson correlations. Because proficiency level is dichotomous, this is a point-biserial correlation and is mathematically equivalent to an independent-samples t-test on the same data: both return the same two-tailed significance value, while r gives the direction and magnitude of the group difference. Reporting r rather than t places the direction and size of every group difference on one standardised scale, comparable across the fifteen categories and with the correlational studies reviewed above. Tables 2 and 3 therefore report r and its two-tailed significance rather than duplicating the t statistic, and this significance value is the one referred to throughout. Because fifteen categories were tested at the .05 level, the two-tailed values were additionally adjusted with the Benjamini–Hochberg procedure (Benjamini & Hochberg, 1995) to control the false discovery rate, and both the uncorrected value and the adjusted q are taken into account before a difference is described as significant. Coefficients are read against Cohen's (1988) benchmarks of .10, .30 and .50 for small, medium and large effects.

To assess whether the items constituting each strategy category could reasonably be aggregated, the internal consistency of every subscale was examined. For the twelve multi-item subscales, Cronbach's alpha ranged from .458 to .793, six of them meeting the conventional .70 criterion (Nunnally & Bernstein, 1994); one improved from .588 to .66 once

a single item with a very low corrected item–total correlation ($r$ = .06) was removed, so that 57 of the 58 items entered the analyses, and no other deletion improved consistency. Because alpha is deflated for very short scales, the three two-item subscales were examined using the inter-item correlation, which ranged from −.04 to .55 and was adequate for one of them. Internal consistency therefore reached an acceptable level for seven of the fifteen categories. All categories were aggregated into category means.

Retaining the weaker subscales was deliberate. The fifteen categories are Nakatani's published taxonomy, not factors derived from these data, and dropping some would make the results incomparable with the studies discussed below, all of which report the inventory in full. Alpha is also a lower bound, deflated by short scales and multidimensionality rather than by noise alone. Because measurement error attenuates observed correlations, the significant differences are conservative estimates, whereas the non-significant ones are ambiguous and are not evidence that the groups behave alike. All category-level comparisons should therefore be read as indicative, and this is acknowledged as a limitation.

## Results and Discussion

### Results

#### *The Use of Communication Strategies Among Different English Proficiencies*

Table 2 presents the use of Speaking strategies.

**Table 2**
*The Use of Speaking Strategies*

| Strategies | Levels | Mean | Std. Deviation | Pearson Correlation | Sig. (2-tailed) |
|---|---|---|---|---|---|
| Social affective strategies | E | 3.30 | .59264 | .445** | < .001 |
| | I | 3.83 | .49020 | | |
| Fluency-oriented strategies | E | 3.59 | .63323 | .164 | .101 |
| | I | 3.79 | .58134 | | |
| Negotiation for meaning while speaking strategies | E | 3.61 | .84860 | .116 | .248 |
| | I | 3.79 | .68699 | | |
| Accuracy-oriented strategies | E | 3.16 | .67701 | .127 | .206 |
| | I | 3.33 | .61035 | | |

| Strategies | Levels | Mean | Std. Deviation | Pearson Correlation | Sig. (2-tailed) |
|---|---|---|---|---|---|
| Message reduction and alteration strategies | E | 4.12 | .66796 | .035 | .728 |
| | I | 4.17 | .68236 | | |
| Non-verbal strategies while speaking | E | 3.59 | 1.09173 | .118 | .240 |
| | I | 3.82 | .86542 | | |
| Message abandonment strategies | E | 3.39 | .82088 | -.359** | < .001 |
| | I | 2.71 | .94000 | | |
| Attempt to think in English strategies | E | 3.87 | .84472 | -.099 | .325 |
| | I | 3.70 | .81025 | | |

*Note.* E = elementary ($n$ = 50); I = intermediate ($n$ = 51). Proficiency level was coded E = 1 and I = 2, so that a positive correlation indicates more frequent use by the intermediate group and a negative correlation more frequent use by the elementary group. Sig. values are two-tailed. * $p < .05$. ** $p < .01$.

As can be seen in Table 2, "Message reduction and alteration strategies" were the most frequently used strategies at both the elementary and the intermediate level ($M$ = 4.12 and 4.17 respectively), and no statistically significant difference was found between the two groups in the use of these strategies (Sig. = .728). "Accuracy-oriented strategies" were the least frequently used by the elementary group ($M$ = 3.16) and the second least frequently used by the intermediate group ($M$ = 3.33), again with no statistically significant difference between them (Sig. = .206). In addition, "Message abandonment strategies" were the least frequently used category among students with higher language ability ($M$ = 2.71). Both group means fall within the "sometimes" band, but the elementary group used these strategies considerably more often ($M$ = 3.39), producing a significant difference between the two groups (Sig. < .001). "Social affective strategies" also showed a significant difference in the frequency of use (Sig. < .001), being more frequently applied by the intermediate group ($M$ = 3.83, often) than by the elementary group ($M$ = 3.30, sometimes). The four remaining strategies were often used by both groups ($M$ = 3.59 to 3.87) with no significant difference (Sig. = .101 to .325).

Table 3 presents the use of Listening strategies.

**Table 3**
*The Use of Listening Strategies*

| Strategies | Levels | Mean | Std. Deviation | Pearson Correlation | Sig. (2-tailed) |
|---|---|---|---|---|---|
| Negotiation for meaning while listening | E | 4.10 | .61614 | -.279** | .005 |
| | I | 3.70 | .77500 | | |
| Fluency-maintaining | E | 3.46 | .71931 | .218* | .029 |
| | I | 3.74 | .57511 | | |
| Scanning | E | 3.82 | .71146 | -.117 | .244 |
| | I | 3.65 | .73276 | | |
| Getting-the-gist | E | 3.58 | .74303 | .013 | .897 |
| | I | 3.59 | .68646 | | |
| Non-verbal while listening | E | 4.02 | .84188 | -.190 | .057 |
| | I | 3.73 | .67218 | | |
| Less-active listener | E | 3.84 | .86396 | -.288** | .003 |
| | I | 3.30 | .94583 | | |
| Word-oriented | E | 3.92 | .57855 | -.086 | .392 |
| | I | 3.81 | .62042 | | |

*Note.* E = elementary ($n$ = 50); I = intermediate ($n$ = 51). Proficiency level was coded E = 1 and I = 2, so that a positive correlation indicates more frequent use by the intermediate group and a negative correlation more frequent use by the elementary group. Sig. values are two-tailed. * $p < .05$. ** $p < .01$.

"Negotiation for meaning while listening" strategies were the most frequently used by elementary students ($M$ = 4.10), significantly more often than by their intermediate counterparts ($M$ = 3.70; Sig. = .005), for whom the most frequently used category was "Word-

oriented" (*M* = 3.81), although the elementary mean for that category was higher still (*M* = 3.92). A further difference, significant before correction for multiple comparisons, was found in "Fluency-maintaining" (Sig. = .029), which intermediate students used more frequently than elementary students (*M* = 3.74 and 3.46, respectively), while the difference in "Non-verbal while listening" fell short of significance (*M* = 4.02 and 3.73; Sig. = .057). The largest difference was found in "Less-active listener" (Sig. = .003), as elementary students often used such strategies (*M* = 3.84) while the intermediate group used them only sometimes (*M* = 3.30). "Scanning", "Getting-the-gist" and "Word-oriented" were used comparatively frequently by both groups (*M* = 3.58 to 3.92) with no significant difference (Sig. = .244, .897 and .392).

### *Correlations Between English Proficiency Levels and Oral Communication Strategies Use*

Tables 2 and 3 report the point-biserial coefficients in full, so only their interpretation is given here. Two of the eight speaking categories and three of the seven listening categories were significantly correlated with proficiency level at the conventional threshold. Positive coefficients indicate more frequent use by the intermediate group and negative coefficients more frequent use by the elementary group. Read against Cohen's (1988) benchmarks, the association approaches large for "Social affective", is medium for "Message abandonment" and approaches medium for "Less-active listener" and "Negotiation for meaning while listening", while it is small for "Fluency-maintaining"; "Non-verbal while listening" (*r* = -.190) fell short of significance (Sig. = .057). After Benjamini–Hochberg adjustment across the fifteen tests, the first four remain significant (*q* = .000 to .018), whereas "Fluency-maintaining" (*q* = .086) falls outside the threshold and is treated as provisional below.

Calculated from the means and standard deviations in Tables 2 and 3, the 95% confidence intervals for the four robust differences are [0.32, 0.74] for "Social affective", [-1.03, -0.33] for "Message abandonment", [-0.90, -0.18] for "Less-active listener" and [-0.68, -0.12] for "Negotiation for meaning while listening", with *d* = 0.98, -0.77, -0.60 and -0.57 respectively. Where a category is described as most or least frequently used, the ranking should be read alongside these intervals, since adjacent categories often differ by a tenth of a scale point or less.

## Discussion

### *Speaking Strategies*

Only two of the eight speaking categories distinguished the groups: "Social affective" strategies, used more often by the intermediate group (*r* = .445), and "Message abandonment" strategies, used more often by the elementary group (*r* = -.359). The other six were statistically indistinguishable (*r* = -.099 to .164, Sig. = .101 to .728), although the intermediate group returned the higher mean in six of the eight. This converges with Ansarin (2003), who reported no significant overall difference between proficiency levels while locating differences in particular strategy types, and it qualifies M. H. Bui (2016): the relationship between proficiency and CS use appears category-specific rather than general.

The advantage in "Social affective" strategies was the largest coefficient recorded in the study and converges with Mirzaei and Heidari (2012), Demir et al. (2018) and Nakatani (2005). Chen (2009) attributed four categories to fluent speakers, namely social affective, fluency-oriented, negotiation for meaning and accuracy-oriented, but only the first separated the groups here (Sig. = .101, .248 and .206 for the other three). Within this proficiency range, the affective and

interpersonal dimension of speaking strategy use appeared to track proficiency more closely than the linguistic-processing dimension did.

The elementary group's heavier use of "Message abandonment" is consistent with Nakatani (2006), La-Ortip (2003), Ansarin (2003) and Bien (1999), and the intermediate mean of 2.71 was the lowest figure recorded for any speaking category. "Message reduction and alteration", however, did not follow: it was the most frequently used category at both levels ($M$ = 4.12 and 4.17) and showed no difference at all ($r$ = .035, Sig. = .728). Chen (2009), Mirzaei and Heidari (2012) and Nakatani (2006) treat reduction and abandonment as a single low-proficiency cluster, yet here they behaved in opposite ways. The result aligns with Jindathai (2016) but stands against T. T. H. Nguyen et al. (2022), who placed message reduction among the least used by Vietnamese tertiary students. Simplifying the message may therefore remain a low-cost repair as proficiency grows, and may not be a marker of low ability.

The four remaining categories were used often by both groups ($M$ = 3.59 to 3.87) with no significant difference. "Attempt to think in English" runs counter to Huang and Van Naerssen (1987), the elementary group scoring numerically higher ($M$ = 3.87 against 3.70; Sig. = .325). "Negotiation for meaning while speaking" reproduces the direction reported by Mirzaei and Heidari (2012) without reaching significance (Sig. = .248), as does "Fluency-oriented" against Demir et al. (2018) and Nakatani (2005) (Sig. = .101). None of these categories therefore showed more than a numerical trend, although the literature predicts a proficiency effect in the productive mode.

***Listening Strategies***

Three of the seven listening categories separated the levels, against two of the eight speaking categories, and the direction is reversed: the elementary group returned the higher mean in five of the seven listening categories and in two of the three significant ones. This echoes La-Ortip (2003), who found a higher overall frequency of CS use among less fluent learners. Since the OCSI measures compensatory behaviour rather than listening ability, the elementary group's heavier reported use is best read as an index of greater difficulty in comprehension rather than of greater skill.

"Negotiation for meaning while listening" was the elementary group's most used category ($M$ = 4.10) and was used significantly more often than by the intermediate group ($M$ = 3.70; $r$ = -.279, Sig. = .005), showing greater effort to request clarification, simplification, slower speech and repetition. This runs against Mirzaei and Heidari (2012) and is not predicted by Ting and Lau (2008), whose Malaysian data showed clarification requests and comprehension checks used at a broadly similar rate by both proficiency groups. The OCSI items in this category ask about compensatory responses to a comprehension failure rather than the collaborative meaning-building measured in speaking, which may explain the divergence. The same construct in the productive mode did not distinguish the groups at all (Sig. = .248), so the direction of the effect depends on the modality in which it is measured.

"Fluency-maintaining" was the only listening category in which the intermediate group's higher mean reached significance ($M$ = 3.74 against 3.46; $r$ = .218, Sig. = .029), and it was the least used category among elementary students, in line with Demir et al. (2018) and Nakatani (2005), though the difference does not survive correction across the fifteen tests, while its speaking counterpart did not differentiate the groups (Sig. = .101). "Non-verbal while listening" was used more often by elementary students ($M$ = 4.02 against 3.73; $r$ = -.190),

though the difference did not reach significance (Sig. = .057); its direction is consistent with Mirzaei and Heidari (2012) and La-Ortip (2003), and "Non-verbal while speaking" likewise showed no difference (Sig. = .240). This may suggest that higher-level students attend to rhythm, intonation and pronunciation while the lower level relies on gesture, eye contact and facial expression to compensate for the language it lacks, although the questionnaire records reported frequency of use rather than attention itself.

The largest listening difference was "Less-active listener" ($r$ = -.288, Sig. = .003), used often by elementary students ($M$ = 3.84) but only sometimes by the intermediate group ($M$ = 3.30). This confirms Huang and Van Naerssen (1987), whose claim about thinking in English was not borne out in the speaking data.

"Scanning", "Getting-the-gist" and "Word-oriented" showed no significant difference ($M$ = 3.58 to 3.92; Sig. = .244, .897 and .392, respectively). The groups were almost identical on "Getting-the-gist" ($M$ = 3.58 and 3.59), unlike Mirzaei and Heidari (2012). The three differences that did emerge all fall on the interactional and compensatory side of listening. The text-processing categories yielded no detectable difference, which, given the ambiguity of the non-significant results, cannot itself be read as evidence that the two groups behave alike.

***Contribution of the Study***

Le et al. (2020) found no significant differences in CS use across academic levels among Vietnamese English majors. Grouping non-English majors by proficiency instead, the present study finds differences in five of the fifteen categories, distributed in a patterned rather than a diffuse way. Three contributions follow.

First, and most securely, the study decouples message reduction from message abandonment. The two are commonly treated as a single avoidance cluster attributed to less proficient learners (Chen, 2009; Mirzaei & Heidari, 2012; Nakatani, 2006). Here they behaved in opposite ways: message abandonment was the second strongest correlate of proficiency in the speaking set ($r$ = -.359, Sig. < .001), while message reduction and alteration was the most frequently used category at both levels and showed essentially no association with proficiency ($r$ = .035, Sig. = .728). The contrast rests on positive observations rather than on an absent difference: a substantial coefficient for abandonment, and for reduction the highest category mean at both levels. This makes it the firmest of the present findings. Abandoning a message and simplifying it therefore appear to be distinct behaviours, and treating them as a single construct obscures the difference.

Second, proficiency was associated with differences in three of the seven listening categories but only two of the eight speaking categories, and in opposite directions: the intermediate group stood out in one speaking category, the elementary group in two of the three significant listening categories. Studies in the reviewed literature report both skills without contrasting the two profiles (Jindathai, 2016; Le et al., 2020; Mirzaei & Heidari, 2012). Negotiation for meaning illustrates the point directly, differentiating the groups in the listening mode (Sig. = .005) but not in the speaking mode (Sig. = .248); since Nakatani's (2006) inventory provides parallel categories in both modes, this asymmetry would be invisible to studies reporting a single combined figure. Because the claim rests partly on absent differences, and unreliable measurement can produce absences, it requires replication before it is treated as established.
Third, the study describes OCSI category use for a population the Vietnamese literature has addressed only indirectly, having examined English majors (T. T. Q. Bui & Intaraprasert, 2012,

2013; Le et al., 2020), a single proficiency level among non-English majors (T. T. Nguyen & Nguyen, 2016; H. C. Nguyen, 2023), or mixed populations without a level-based contrast (T. T. H. Nguyen et al., 2022). Social affective and fluency-maintaining strategies are plausible instructional targets at elementary level, while the persistence of message reduction at intermediate level suggests that higher-proficiency learners also need support in expanding rather than simplifying their intended messages.

These contributions should be read alongside the reliability limitations reported under Instrument and Data Analysis. Internal consistency was acceptable for only seven of the fifteen categories, so every category-level comparison is indicative rather than conclusive. The differences that were detected stand on firmer ground than those that were not, since attenuation works against them.

## Limitations and Future Research

Several limitations should be borne in mind when interpreting these findings. First, the study was conducted at a single language centre in one university, with 101 participants drawn from two proficiency levels, so the results cannot be generalised to Vietnamese non-English majors as a whole. Second, participants were selected purposively rather than randomly: students were assigned to levels by the Center's placement test and were then identified by their class teachers as clearly matching the CEFR descriptors for that level. This procedure was adopted to ensure that the two groups were genuinely distinct in proficiency, but it means that the sample is not representative of all students at each level. No formal check of agreement between teachers was carried out, so the level assignments retain an element of subjective judgement. Third, proficiency was treated as a two-level category rather than as a continuous variable, which limits what the correlational analysis can show. A standardised test score would allow the relationship between proficiency and strategy use to be modelled more precisely. Fourth, demographic variables such as gender and year of study were not recorded, so the differences reported here cannot be separated from any imbalance between the two groups on those variables.

Further limitations concern the instrument and the data. The study relied entirely on self-report: students indicated how frequently they believed they used each strategy, which may not correspond to their actual behaviour. As reported under Instrument and Data Analysis, internal consistency was acceptable for only seven of the fifteen categories and cannot be attached to any individual row of Tables 2 and 3, so establishing the reliability of each category separately, on a larger sample, is a necessary step before these differences are treated as settled. Future research could include advanced-level learners, combine the questionnaire with recorded speaking and listening tasks or stimulated-recall interviews, and draw on larger samples across several institutions. Intervention studies on explicit strategy training would also be valuable.

## Conclusion

This study compared how elementary and intermediate non-English majors use oral communication strategies. Of the eight speaking categories only two differentiated the groups: intermediate students used "Social affective" strategies significantly more often and elementary students used "Message abandonment" significantly more often, while "Message reduction and alteration" was the most frequently used category at both levels with no significant difference. Of the seven listening categories, three differentiated the groups: elementary students used "Negotiation for meaning while listening" and "Less-active listener"

more often, whereas intermediate students used “Fluency-maintaining” more often. Proficiency was thus associated with more apparent group differences in listening than in speaking, although the numerical asymmetry rests on a listening difference that did not survive adjustment for multiple testing, and among the listening categories the difference lay predominantly with the elementary group. For teaching, CSs should be introduced explicitly and practised in class (Dörnyei, 1995), with the strategies targeted chosen according to what learners at each level already do and under-use.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that Claude, an AI assistant developed by Anthropic, was used in preparing this manuscript to check the consistency of the reported results and references, to draft and revise parts of the discussion, and to condense, edit, and format the text. The research design, data collection, and analysis are the authors’ own. The authors reviewed and approved all text and take full responsibility for the content of the manuscript.

## References

Ansarin, A. A. (2003). Still a new inventory of communication strategies? The case of Iranian EFL learners. *Journal of the Faculty of Letters and Humanities of Tabriz University, 46*(187), 17–43.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B, 57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bialystok, E. (1990). *Communication strategies: A psychological analysis of second language use*. Basil Blackwell.

Bien, T. K. (1999). *An investigation into the use of specific communication strategies used by adult Vietnamese learners* [Unpublished Master's thesis]. University of South Australia.

Bui, M. H. (2016). *Exploring English oral communication strategies used by Vietnamese students in transnational education advanced programs in Vietnam* (Publication No. 9921863825601341) [Doctoral dissertation, RMIT University]. https://researchrepository.rmit.edu.au/esploro/outputs/9921863825601341

Bui, T. T. Q., & Intaraprasert, C. (2012). Gender, high school background and use of strategies by English majors in Vietnam for coping with communication breakdowns. *International Journal of Scientific and Research Publications, 2*(12), 1–17.

Bui, T. T. Q., & Intaraprasert, C. (2013). The effects of attitude towards speaking English and exposure to oral communication in English on use of communication strategies by English majors in Vietnam. *International Journal of Scientific and Research Publications, 3*(2), 1–9.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 2–27). Longman.

Chen, H. W. (2009). *Oral communication strategies used by English major college students in Taiwan* [Unpublished master's thesis]. Chaoyang University of Technology.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Corder, S. P. (1983). Strategies of communication. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 15–19). Longman.

Demir, Y., Mutlu, G., & Şişman, Y. S. (2018). Exploring the oral communication strategies used by Turkish EFL learners: A mixed methods study. *International Journal of Instruction, 11*(2), 539–554. https://doi.org/10.12973/iji.2018.11237a

Dörnyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly, 29*(1), 55–85. https://doi.org/10.2307/3587805

Faerch, C., & Kasper, G. (1980). Processes and strategies in foreign language learning and communication. *Interlanguage Studies Bulletin Utrecht, 5*(1), 47–118. http://files.eric.ed.gov/fulltext/ED213247.pdf

Huang, X.-H., & Van Naerssen, M. (1987). Learning strategies for oral communication. *Applied Linguistics, 8*(3), 287–307. https://doi.org/10.1093/applin/8.3.287

Jindathai, S. (2016). Use of communication strategies by engineering students in relation to exposure to oral communication in English at a Thai private higher education institution. *LEARN Journal: Language Education and Acquisition Research Network*, *9*(2), 135–154. https://so04.tci-thaijo.org/index.php/LEARN/article/view/102652

La-Ortip, W. (2003). *English language studies: Oral communication strategies employed by English majors taking listening and speaking 1 at Rajabhat Institute Nakhon Ratchasima, Thailand* [Master's thesis, Suranaree University of Technology]. http://sutir.sut.ac.th:8080/jspui/handle/123456789/1138

Le, V. T., Huynh, T. A., & Tran, K. H. (2020). Strategies used by undergraduate English-majored students in oral communication. *VNU Journal of Foreign Studies*, *36*(1), 156–179. https://doi.org/10.25073/2525-2445/vnufs.4506

Mirzaei, A., & Heidari, N. (2012). Exploring the use of oral communication strategies by (non)fluent L2 speakers. *The Journal of Asia TEFL*, *9*(3), 131–156. http://www.asiatefl.org/main/download_pdf.php?i=92&c=1419299763

Mitchell, R., & Myles, F. (1998). *Second language learning theories*. Arnold.

Nakatani, Y. (2005). The effects of awareness-raising training on oral communication strategy use. *The Modern Language Journal*, *89*(1), 76–91. https://doi.org/10.1111/j.0026-7902.2005.00266.x

Nakatani, Y. (2006). Developing an oral communication strategy inventory. *The Modern Language Journal*, *90*(2), 151–168. https://doi.org/10.1111/j.1540-4781.2006.00390.x

Nguyen, H. C. (2023). Enhancing EFL Vietnamese students' communicative competence through selected speaking strategies. *REiLA: Journal of Research and Innovation in Language*, *5*(1), 17–31. https://doi.org/10.31849/reila.v5i1.12276

Nguyen, T. T., & Nguyen, T. K. T. (2016). Oral English communication strategies among Vietnamese non-majors of English at intermediate level. *American Journal of Educational Research*, *4*(3), 283–287. https://doi.org/10.12691/education-4-3-9

Nguyen, T. T. H., Vu, V. P., Yen, T. T., & Do, T. T. Y. (2022). The influences of oral communication strategies on students' English speaking proficiency. *Can Tho University Journal of Science, 14*(2), 8–17. https://doi.org/10.22144/ctu.jen.2022.012

Nugroho, A. P. (2019). Communication strategies used by EFL learners with different English achievements in oral communication. *Indonesian Journal of Language Teaching and Linguistics*, *4*(3), 138–155. https://doi.org/10.30957/ijoltl.v4i3.606

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Oweis, T. (2013). A literature review on communication strategies in language learning. *European Scientific Journal*, *9*(26), 245–267. https://eujournal.org/index.php/esj/article/view/1799

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Newbury House.

Phuong, H. Y., & Tran, T. T. Q. (2021). English teaching reform in Vietnam: Responses of non-English majored students toward the target English level of the CEFR-V. *Journal of Language and Linguistic Studies*, *17*(2), 1189–1204. https://doi.org/10.52462/jlls.86

Rabab'ah, G. (2003). Communication problems facing Arab learners of English. *Journal of Language and Learning, 3*(1), 180–197. https://files.eric.ed.gov/fulltext/ED473079.pdf

Rabab'ah, G. (2016). The effect of communication strategy training on the development of EFL learners' strategic competence and oral communicative ability. *Journal of Psycholinguistic Research, 45*(3), 625–651. https://doi.org/10.1007/s10936-015-9365-3

Tarone, E. (1980). Communication strategies, foreigner talk, and repair in interlanguage. *Language Learning, 30*(2), 417–431. https://doi.org/10.1111/j.1467-1770.1980.tb00326.x

Tarone, E. (1981). Some thoughts on the notion of communication strategy. *TESOL Quarterly, 15*(3), 285–295. https://doi.org/10.2307/3586754

Ting, S. H., & Lau, L. Y. (2008). Lexical and discourse-based communication strategies of Malaysian ESL learners. *Malaysian Journal of ELT Research*, *4*, 18–31. https://melta.org.my/journals/MAJER/downloads/majer04_01_02.pdf

Tran, T. T. (2013). Factors affecting teaching and learning English in Vietnamese universities. *The Internet Journal Language, Culture and Society*, *38*, 138–143. http://aaref.com.au/en/publications/journal/

Trinh, T. T. H., & Mai, T. L. (2018). Current challenges in the teaching of tertiary English in Vietnam. In J. Albright (Ed.), *English tertiary education in Vietnam* (1st ed., pp. 40–53). Routledge. https://doi.org/10.4324/9781315212098-4

Tuan, N. H., & Mai, T. N. (2015). Factors affecting students' speaking performance at Le Thanh Hien high school. *Asian Journal of Educational Research*, *3*(2), 8–23. https://www.multidisciplinaryjournals.com/wp-content/uploads/2015/03/FACTORS-AFFECTING-STUDENTS%E2%80%99-SPEAKING.pdf

**Contact email:** vohuyennhi@tdtu.edu.vn
nguyenhuutai@tdtu.edu.vn

**ORCID:** https://orcid.org/0009-0008-5168-5616
https://orcid.org/0009-0006-8490-7233

# Safety Experience and Student Well-being at School: The Moderating Role of School Characteristics Among Indonesian Students

Septian Cahyo, Center for Educational Assessment, Indonesia
Siti Raudlah, Center for Educational Assessment, Indonesia
Januar Pribadi, Center for Educational Assessment, Indonesia

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

Well-being has become an important factor that contributes to students' mental health. Since prior studies suggested that a safe school environment significantly increases students' well-being, this study aims to investigate the effects of student's safety experience at school, reflected by being safe from bullying, sexual abuse, and physical punishment, on their perceived well-being, and how these relationships are moderated by specific school characteristics that includes geographic location, school status, and school type. Using data from 27,154 high schools in the 2024 Indonesian National Assessment, the results showed that safety experience was positively associated with students' well-being, where safety from bullying became the strongest predictor, followed by safety from physical punishment and sexual abuse. These results suggest that the safer a school environment is, the higher the quality of students' well-being becomes. A follow-up analysis using multiple regression with interaction terms revealed that the effects of being safe from bullying on well-being were significantly different across school characteristics, where the impacts were stronger for urban and public schools compared with rural and private schools, but weaker for vocational compared with general high schools. While the effects of sexual abuse on well-being was significantly moderated by school status, the effects of safety from physical punishment varied significantly across geographic location and school type. The theoretical and practical implications of the findings are discussed, along with the limitations and suggestions for future studies.

*Keywords:* well-being, safety experiences, bullying, sexual abuse, physical punishment, moderation effects, school characteristics

iafor
The International Academic Forum
www.iafor.org

## Introduction

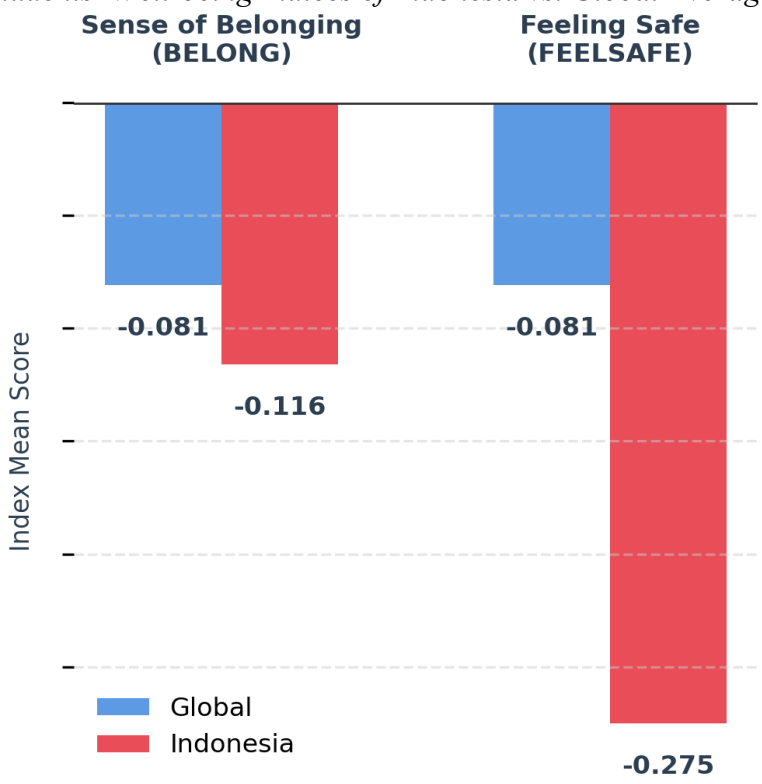
Well-being has become a fundamental factor that contributes to students' quality of life. Previous research has shown that promoting students' well-being not only prevents mental disorders such as depression and anxiety (Schotanus-Dijkstra et. al., 2017), but also empowers them to feel connected to school, have a purpose, and enjoy learning activities (Arslan & Renshaw, 2018) that leads to enhanced academic outcomes (Kaya & Erdem, 2021).

As a part of educational reform, as mandated in the Ministerial Regulation (*Permendikdasmen*) Number 6 of 2026, Indonesia has shifted its education paradigms from merely focus on academic achievements to a more holistic focus on students' well-being and schools' safety. As a part of this spirit, the country has transformed its curriculum and national examination system that previously captured only cognitive aspects with a new national assessment instrument that accommodates students' basic competencies, their non-cognitive skills, and the safety of the learning environment at schools.

While many policies have been implemented to support schools' safety environment, Indonesia faces challenges promoting students' quality of life. According to the Programme for International Student Assessment in 2022, the well-being level of Indonesian students, reflected by their perceived sense of belonging (BELONG) and feeling safe (FEELSAFE) at schools, are still relatively low compared with the global average (OECD, 2022). As depicted in Figure 1, while the global average of students' sense of belonging and their perceived safety feeling at schools are considerably low at -0.081 points, Indonesia saw more concerning levels at -0.116 and -0.275 for both variables consecutively. This raises a concern that the country is struggling when it comes to improving students' well-being.

**Figure 1**

*Students' Well-being Indices of Indonesia vs. Global Average*

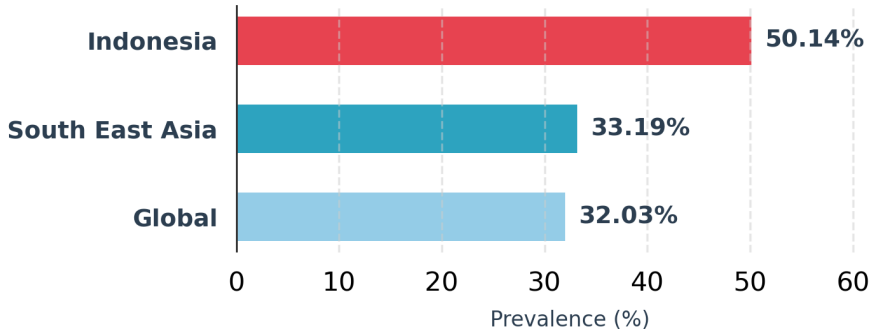


## Bullying and Well-being

Research has consistently shown that bullying has detrimental effects on students' well-being. Arslan et al. (2021) conducted a study toward adolescence in Turkey and found that students who engage in bullying activities are more prone to experience emotional and behavioral problems. Another study by Hellfeldt et al. (2020) revealed that experiencing cyberbullying significantly increases depressive and anxiety symptoms and decreases subjective well-being. Surprisingly, both studies suggested that the impacts apply not only to the victims but also to the perpetrators.

According to the Global School-based Student Health Survey (GSHS) between 2003 and 2015, 32.03% of adolescents aged 12–17 years in 65 countries experienced bullying in the past 30 days when the survey was conducted (Man et al., 2022). As presented in Figure 2, the prevalence was slightly higher in South East Asia where 33.19% of the students in this region suffered from bullying activities. Ironically, the prevalence was considerably higher for Indonesia where 50.14% of the students reported have become the victims of bullying. These conditions make us believe that it is important to study the effects of bullying on students' well-being in the Indonesian context.

**Figure 2**
*The Prevalence of Bullying in Indonesia, South East Asia, and Global*




## Sexual Abuse

Few studies examine how sexual harassment affects students' well-being while considering sociodemographic factors and possible mediating or moderating mechanisms. A study by Bendixen et al. (2018) among 2,869 Norwegian students in 2007 and 2013/2014, showed that students who experienced more non-physical sexual harassment reported higher levels of anxiety and depression, lower self-esteem, and more negative body image. Further, they suggested that harassment remained a significant predictor of poorer well-being even after controlling for gender, age, vocational education, parents' status, sexual minority status, immigrant status, sexual coercion, and prior forced intercourse.

Similarly, using a population-based survey among 2,873 South Australian adults in 2015, González-Chica et al. (2019) found that bullying and sexual abuse or harassment-related experiences were associated with harmful behaviours, antidepressant use, and lower health-related quality of life. Although the study focused on adults, it is relevant to our research because it shows the longer-term association between experiences of bullying or sexual abuse and well-being-related outcomes. Given the limited research on sexual abuse and students' well-being in Indonesian context, and considering the long-term impacts throughout the victims future life, it is crucial to investigate these associations among Indonesian students.

## Physical Punishment

Physical punishment, or usually referred to as corporal punishment in several research studies, becomes an important part of Indonesian safety schools' climate. While studies about this were very limited in Indonesia, the World Health Organization (2025) stated that this action remains common in many schools globally. For example, based on a meta-analysis studies between 1980 and 2017 in Africa and Central America, 70% of students suffered from experiencing physical punishment at schools.

Additionally, Bangash et al. (2024) conducted a study among 120 male Pakistani students from grades 9 and 10 who had experienced corporal punishment three to five times within an academic year. The results suggested that experiencing corporal punishment leads to increased fear, anxiety and aggression as well as decreased self-confidence that impacts students well-being. Furthermore, they also found that corporal punishment trigger students to experience suicidal thought. Although some limitations of their study were mentioned, this suggests that further investigation is needed to better understand the mechanisms of physical punishment on well-being.

## Schools' Characteristics

While many studies showed the direct effects of bullying, sexual abuse, and physical punishment on students well-being, some scholars believed that it is not only influenced by direct factors, but may also involve underlying mechanisms or conditions that strengthen or weaken these relationships (Freire et al., 2019; Phan et al., 2016). This motivates us to further extend our interest not only on the direct effects of a student's safety experience at school but also the moderating roles of some schools' characteristics to unveil the interaction effects.

Sharma et al. (2025) found that there were significant differences in students' well-being levels based on gender and educational qualification. Meanwhile, demographic factors such as residential background and age group did not show significant differences in the well-being levels of students in Haryana, India. Since Indonesia and India may have different cultural and socioeconomic structure, it is worth replicating the studies to investigate if the results may be different in terms of geographic location (which are urban and rural areas).

Beside the geographic location, the school status (public or private) can be an interesting moderator since they both have different characteristics either from management or resources. In 2019, Dos Santos et al. conducted a study among 2,280 students aged 9 to 13 years from Rio Grande do Sul state in Brazil and found that students who lived in rural and studied in public schools, and those who lived in urban and studied in private schools, have higher average levels of subjective well-being. While their study examined the interaction effects across geographic location and school status, it is interesting to treat school status as a moderator effect of schools' safety climate on students' well-being in the Indonesian context.

Finally, the last moderating factor that becomes our interest is school type which are general or vocational highschool. While it was challenging to dig on the literature that investigates well-being at the vocational high school, we found a study from Mewafarosh and Agarwal (2022) that demonstrated the significant differences between most demographic variables and subjective well-being levels of students who enrolled in a premier business school. Since vocational high school has unique and important dynamics in the Indonesian education system, this limited literature motivates us to include this variable as one of our moderators.

## Research Questions

Based on our exploration of previous studies, and considering the importance of these factors for student well-being, our research dives into these two big questions:

1. To what extent do the student's safety experience from bullying, sexual abuse, and physical punishment predict their perceived well-being at school?
2. How are these relationships moderated by geographic location (urban vs. rural), school status (public vs. private), and school type (vocational vs. general high school)?

# Methodology

## Data Collection

This study utilizes data from Indonesian National Assessment, specifically Learning Environment Survey (*Survey Lingkungan Belajar, Sulingjar*) 2024. The survey was conducted on August 19th-22nd, 2024 across 27.154 High Schools (General of Vocational) in Indonesia. At each school, 45 students in grade 11 were randomly selected using stratified sampling, taking into account the proportion of students by class and gender. Student individual scores were aggregated at the school levels and became the unit of analysis for this study. The data covers 38 Provinces that consist of 513 Districts (98 Cities and 415 Regencies).

## Variables

The variables and how they are measured are briefly explained as follows:

### *Well-being*

Students' well-being was measured using several item questionnaires that asked students to rate their perceived well-being levels and experience. For example, "How often have you experienced the following at school?: During break time or physical education classes, I worry that my belongings might be stolen at school". Students then give a response using Likert scales such as: "Never"; "Rarely"; "Often"; "Always"; and "Prefer not to answer". After being aggregated at the school level, the scale ranges from 20.87 to 97.58 with a mean of 65.06 (SE = 0.05) and standard deviation of 8.27. The higher score of well-being indicates students experience a more positive level of well-being.

### *Safety From Bullying*

Bullying was measured using several item questionnaires that asked students' experience of bullying activities at school. For example, "During the past year, how often have other students or your teachers engaged in the following behaviors?: Called you or your friends by names that highlight a perceived weakness (e.g., 'stupid', 'troublemaker', or similar names)". Students then give a response using Likert scales such as: "Never"; "Rarely (fewer than 5 times in the past year)"; "Often (5–10 times in the past year)"; "Very often (more than 10 times in the past year)"; and "Prefer not to answer/Not applicable". The scale ranges from 1.29 to 100 with a mean of 68.12 (SE = 0.05) and standard deviation of 8.37. Since the index measures schools' safety experience, a higher score of this variable means that the school environment is safer from bullying activities.

### *Safety From Sexual Abuse*

Sexual abuse was measured using several item questionnaires that asked students' experience related to sexual abuse at school. For example, "How often do you experience the following at school?: Another student or an adult touched parts of your body that are normally covered by clothing". Students then give a response using Likert scales such as: "Never"; "Rarely"; "Often"; "Prefer not to answer"; and "Not applicable". The scale ranges from 0 to 95.82 with a mean of 80.16 (SE = 0.05) and standard deviation of 8.56. Similar to bullying, a higher score for this variable means that the school environment is safer from sexual abuse.

### *Physical Punishment*

Physical punishment was measured using several item questionnaires that asked students' experience in receiving physical punishment at school. For example, "When you arrive late to school, have you ever received any of the following punishments?: Run laps around the school field". Students then give a response using likert scales such as: "Never"; "Rarely"; "Often"; "Always", and "Prefer not to answer". The scale ranges from 0 to 100 with a mean of 74.73 (SE = 0.08) and standard deviation of 12.37. A higher score means that the school is safer from physical punishment.

### *Schools' Characteristics*

**Geographic Location.** This is a categorical variable with two levels: urban and rural. The variable was derived from the country's district information where Cities are translated as urban while Regencies as rural. Since the unit of analysis is at the school level, there are 6,384 (23.51%) schools located in urban areas and 20,770 (76.49%) schools located in rural areas.

**Schools' Status.** This is also a categorical variable that indicates the schools' source of funding whether they are funded by the government (Public) or independent groups or organizations (Private). Among 27.154 high schools, 10,528 (38.77%) of them are Public Schools and the rest 16,626 (61.23%) are Private Schools.

**Schools' Type.** This variable categorizes a school into General (*SMA*) or Vocational (*SMK*) High School. The proportion of Indonesian schools by type are relatively balanced at 13,862 (51.05%) and 13,292 (48.95%) for General and Vocational High Schools, respectively.

The complete summary of the variables, including the mean of each continuous variable at each level of categorical variable, are summarized in Table 1.

**Table 1**
*Summary of the Variables*

| Variable Names | Min | Max | Mean | SE | SD |
|---|---|---|---|---|---|
| Continuous Variables | | | | | |
| *Well-being* | 20.87 | 97.58 | 65.06 | 0.05 | 8.27 |
| *Safety from Bullying* | 1.29 | 100.00 | 68.12 | 0.05 | 8.37 |
| *Safety from Sexual Abuse* | 0.00 | 95.82 | 80.16 | 0.05 | 8.56 |
| *Safety from Physical Punishment* | 0.00 | 100.00 | 74.73 | 0.08 | 12.37 |
| | **N (%)** | **WB** | **BUL** | **XAB** | **PHY** |
| Geographic Location | | | | | |
| *Urban* | 6,384 (23.51%) | 66.50 | 70.44 | 82.73 | 79.20 |
| *Rural* | 20,770 (76.49%) | 64.62 | 67.41 | 79.37 | 73.35 |
| Schools' Status | | | | | |
| *Public* | 10,528 (38.77%) | 66.50 | 69.33 | 81.47 | 76.15 |
| *Private* | 16,626 (61.23%) | 64.15 | 67.36 | 79.33 | 73.83 |
| Schools' Type | | | | | |
| *General High School* | 13,862 (51.05%) | 67.02 | 69.04 | 81.21 | 76.46 |
| *Vocational High School* | 13,292 (48.95%) | 63.02 | 67.17 | 79.06 | 72.92 |

*Note.* SE = Standard Error of The Mean; SD = Standard Deviation; N = number of schools
WB = Well-being; BUL = Bullying; XAB = Sexual Abuse; PHY = Physical Punishment

## Data Analysis

To answer our research questions, we will utilize the multiple regression analysis using the Ordinary Least Square (OLS) estimation. For our first research question, our regression model is expressed as follows:

Model 1. $$\widehat{Wellbeing}_i = \alpha + \beta_1(urban)_i + \beta_2(public)_i + \beta_3(vocat)_i + \beta_4(bullying)_i + \beta_5(sexabs)_i + \beta_6(punish)_i + \varepsilon_i$$

For our second research question, the multiple regression analysis with interaction terms is conducted, expressed as follows:

Model 2. $$\widehat{Wellbeing}_i = \alpha + \beta_1(urban)_i + \beta_2(public)_i + \beta_3(vocat)_i + \beta_4(bullying)_i + \beta_5(sexabs)_i + \beta_6(punish)_i + \beta_7(urban{:}\,bullying)_i + \beta_8(urban{:}\,sexabs)_i + \beta_9(urban{:}\,punish)_i + \beta_{10}(public{:}\,bullying)_i + \beta_{11}(public{:}\,sexabs)_i + \beta_{12}(public{:}\,punish)_i + \beta_{13}(vocat{:}\,bullying)_i + \beta_{14}(vocat{:}\,sexabs)_i + \beta_{15}(vocat{:}\,punish)_i + \varepsilon_i$$

In addition, we also conduct an incremental F-test to see if adding interaction terms in the Model 2 yields a better fit for model prediction.

## Results

Table 2 summarizes the findings of our regression models. Model 1 investigates the main effects of schools' safety environment on well-being after controlling for school characteristics (Research Question 1), while Model 2 compares the predictors by accounting for interactions between each domain of the safety environment and school characteristics (Research Question 2).

**Table 2**
*Summary of Regression Estimate*

| Predictors | B | Beta (β) | SE | t | p-value | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower Limit | Upper Limit |
| **Model 1**[a] | | | | | | | |
| Intercept | 32.135 | - | 0.422 | 76.139 | < 0.001* | 31.308 | 32.962 |
| urban | 0.183 | 0.009 | 0.103 | 1.766 | 0.077 | -0.020 | 0.386 |
| public | 0.751 | 0.044 | 0.091 | 8.249 | < 0.001* | 0.573 | 0.930 |
| vocat | -2.713 | -0.164 | 0.087 | -31.017 | < 0.001* | -2.884 | -2.541 |
| bullying | 0.236 | 0.239 | 0.008 | 28.719 | < 0.001* | 0.220 | 0.253 |
| sexabs | 0.108 | 0.112 | 0.008 | 14.331 | < 0.001* | 0.093 | 0.123 |
| punish | 0.122 | 0.183 | 0.005 | 23.274 | < 0.001* | 0.112 | 0.133 |
| **Model 2**[b] | | | | | | | |
| Intercept | 31.442 | - | 0.721 | 43.631 | < 0.001* | 30.029 | 32.854 |
| urban | -11.201 | -0.574 | 1.190 | -9.410 | < 0.001* | -13.534 | -8.868 |
| public | -9.712 | -0.572 | 0.988 | -9.834 | < 0.001* | -11.647 | -7.776 |
| vocat | 6.138 | 0.371 | 0.834 | 7.356 | < 0.001* | 4.502 | 7.773 |
| bullying | 0.207 | 0.209 | 0.014 | 14.409 | < 0.001* | 0.179 | 0.235 |
| sexabs | 0.132 | 0.137 | 0.013 | 9.834 | < 0.001* | 0.106 | 0.158 |
| punish | 0.131 | 0.196 | 0.009 | 13.813 | < 0.001* | 0.112 | 0.149 |
| urban:bullying | 0.104 | 0.380 | 0.020 | 5.241 | < 0.001* | 0.065 | 0.143 |
| urban:sexabs | 0.027 | 0.114 | 0.019 | 1.429 | 0.153 | -0.010 | 0.064 |
| urban:punish | 0.024 | 0.099 | 0.012 | 2.008 | 0.045* | 0.001 | 0.048 |
| public:bullying | 0.203 | 0.835 | 0.019 | 10.473 | < 0.001* | 0.165 | 0.241 |
| public:sexabs | -0.052 | -0.251 | 0.018 | -2.893 | 0.004* | -0.087 | -0.017 |
| public:punish | 0.009 | 0.042 | 0.012 | 0.797 | 0.425 | -0.013 | 0.032 |
| vocat:bullying | -0.058 | -0.240 | 0.017 | -3.526 | < 0.001* | -0.091 | -0.026 |
| vocat:sexabs | -0.008 | -0.041 | 0.015 | -0.554 | 0.580 | -0.038 | 0.021 |
| vocat:punish | -0.056 | -0.252 | 0.011 | -5.264 | <0.001* | -0.076 | -0.035 |

[a]Multiple $R^2$= 0.294; Adjusted $R^2$= 0.294; $F(6, 27147) = 1885.00$, *p-value* < 0.001

[b]Multiple $R^2$= 0.309; Adjusted $R^2$= 0.309; $F(15, 27138) = 809.20$, *p-value* < 0.001

*Note.* B = Unstandardized Coefficient; Beta (β) = Standardized Coefficient; SE = Standard Error

## Direct Effects of School Safety Experience on Well-being

This section is aimed to answer our first research question. According to Table 2, Model 1 yielded 0.294 of adjusted $R^2$. This means that our model explained 29.4% of the variation of students' well-being. The model was statistically significant at $F(6, 27147) = 1885.00$, $p < 0.001$. After controlling for schools' characteristics, the result suggests that all domains of safety experience were significant predictors of well-being ($t < 0.001$).

Unstandardized coefficient expresses the raw magnitude of how much well-being increases or decreases by any changes of the predictors. Since all of the coefficients were positive, it suggests that the more students feel safe from bullying, sexual abuse, or physical punishment, the higher their well-being is. For example, for every one unit increase in safety from bullying, the students' well-being increases by 0.236, on average, holding constant other variables.

The standardized coefficient (β) expresses the magnitude of how much well-being increases or decreases by any changes of the predictors in standardized score. Since they are already in the same scale, we can conveniently compare which predictors contribute the strongest effects on a dependent variable. In this case, safety from bullying seemed to be the strongest predictors of well-being because for every one standard deviation increase in safety from bullying is associated with an increase of average well-being by 0.239 standard deviation, holding constant other variables.

## The Moderating Role of Schools Characteristics

For our second research question, we found some significant interaction terms that moderates the effects of school safety domains on well-being. As presented in Table 2 (Model 2), for safety from bullying, its effects on well-being were significantly moderated by geographic location, school status, and school type. A positive coefficient in urban: bullying suggests that the effects of safety from bullying on well-being is higher for students who attend urban schools compared to rural schools. As an illustration, holding constant other variables, a one-unit increase of safety from bullying for a student in urban school is associated with an average increase of well-being by 0.311 points (0.207 + 0.104). Meanwhile, the average increase in well-being for every one-unit increase of safety from bullying is only 0.207 for a student from rural school. Additionally, the results also indicate that the effects of safety from bullying on well-being were significantly higher for public schools compared with private schools and lower for vocational high school compared with general high school.

In terms of safety from sexual abuse, we found no significant moderator effects regarding geographic location and school type, suggesting that the impacts of sexual abuse are uniform among these school characteristics. However, we found that the effect was significantly different among school status where the slope is slightly steeper for private schools compared with public schools. For safety from physical punishment, the effects on well-being did not significantly vary across school status. However, we found that the effects were significantly higher in urban schools and lower in vocational schools.

The model with interaction effects (Model 2) showed a higher performance by explaining 30.9% variance of well-being (Adjusted $R^2 = 0.309$) compared with a model without interactions (Model 1). To check if the $R^2$ change was meaningful, we conducted a model comparison using incremental F-test. As shown in Table 3, although the improvement of

variance explained was only 1.5%, the difference was statistically significant, suggesting that Model 2 has a better fit in predicting students well-being compared with Model 1.

**Table 3**
*Model Comparison Using Incremental F-test*

| Model | $R^2$ | $\Delta R^2$ | F | df1, df2 | ΔF | Δdf | p-value |
|---|---|---|---|---|---|---|---|
| Model 1 (Main Effects) | 0.294 | - | 1884.588 | 6, 27147 | - | - | - |
| Model 2 (Interactions) | 0.309 | 0.015 | 809.156 | 15, 27138 | 65.384 | 9 | < 0.001 |

*Note.* $\Delta R^2$ = Change in $R^2$; ΔF = Change in F Statistics; Δdf = Change in Degree of Freedom

## Discussion

This study aims to investigate the effects of safety experience at schools on student well-being, including the moderating role of school characteristics. In terms of the main effects, our findings suggests that safety from bullying, sexual abuse, and physical punishment were significant predictors of well-being where bullying becomes the strongest predictor, followed by physical punishment and sexual abuse. This result is consistent with our literature review, especially bullying, where it has strong correlation with anxiety and depressive symptoms and behavioral problems. Although sexual abuse appeared to be the weakest predictor of well-being, it does not mean that it is less important. Since this is considered a sensitive issue in Indonesian society, the prevalence of these variables may be undermined compared to bullying or physical punishment that are more common and less sensitive.

Since bullying became the strongest predictor, it is followed by significant interaction effects across all school characteristics domains, indicating that the relationship between bullying and students' well-being is context dependent. Therefore, it requires localized interventions and mitigation at each level of school characteristics.

No significant interaction effects were found between sexual abuse and well-being across geographic location and school type, suggests that the impacts of sexual abuse are uniform and therefore requires full attentions in every level of school characteristics domain.

### Implications

This study contributes to the development of theoretical understanding of bullying, sexual harassment or sexual abuse, and student well-being in the Indonesian context. The findings may help explain whether existing theories developed in other countries are also relevant for understanding students' well-being in Indonesia, especially when considering school-level and contextual factors.

### Limitations

Our study used a cross-sectional design, meaning that the independent and dependent variables were measured at the same time. Consequently, the findings can only be interpreted as associations, rather than causal relationships. In addition, our study utilized self-reported data which may be prone to bias, such as social desirability bias, recall bias, underreporting, or overreporting. Moreover, the observation of our data was aggregated at the school level, so the study could not capture variation at the individual student level. This limits the ability to explain how bullying, sexual harassment, or well-being differ among students within the same school.

## Suggestions

For future studies, it is worthwhile to include additional student-level variables that are relevant to enrich our understanding about well-being. This includes gender, socioeconomic status, grade level, school type, or family background. We also suggest applying multilevel analysis to examine both student-level and school-level factors to help distinguish whether differences in well-being are explained by individual characteristics, school characteristics, or both.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Gemini was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The author further declares that, apart from Gemini, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

Arslan, G., Allen, K. A., & Tanhan, A. (2021). School bullying, mental health, and wellbeing in adolescents: Mediating impact of positive psychological orientations. *Child Indicators Research, 14*(3), 1007–1026.

Arslan, G., & Renshaw, T. L. (2018). Student subjective wellbeing as a predictor of adolescent problem behaviors: A comparison of first-order and second-order factor effects. *Child Indicators Research, 11*(2), 507–521.

Bangash, M., Nayab, T., Zubair, M., Nawaz, H., Ullah, I., & Ullah, S. (2024). Examining the impacts of corporal punishment in secondary schools: Unraveling psychological repercussions on students. *Journal of Health and Rehabilitation Research, 4*(2), 364-369.

Bendixen, M., Daveronis, J., & Kennair, L.E.O. (2018). The effects of non-physical peer sexual harassment on high school students' psychological well-being in Norway: consistent and stable findings across studies. *Int J Public Health, 63*, 3–11. https://doi.org/10.1007/s00038-017-1049-3

Dos Santos, B. R., Sarriera, J.C., & Bedin, L.M. (2019). Subjective Well-Being, Life Satisfaction and Interpersonal Relationships Associated to Socio-Demographic and Contextual Variables. *Applied Research Quality Life, 14*, 819–835. https://doi.org/10.1007/s11482-018-9611-6

Freire, C., Ferradás, M. D. M., Núñez, J. C., Valle, A., & Vallejo, G. (2019). Eudaimonic well-being and coping with stress in university students: The mediating/moderating role of self-efficacy. *International journal of environmental research and public health, 16*(1), 48. https://doi.org/10.3390/ijerph16010048

González-Chica, D. A., Licinio, J., Musker, M., Wong, M., Bowden, J., Hay, P., Chittleborough, C., & Stocks, N. (2019). Bullying and sexual abuse and their association with harmful behaviours, antidepressant use and health-related quality of life in adulthood: a population-based study in South Australia. *BMC public health, 19*(1), 26. https://doi.org/10.1186/s12889-018-6367-8

Hellfeldt, K., López-Romero, L., & Andershed, H. (2020). Cyberbullying and psychological well-being in young adolescence: the potential protective mediation effects of social support from family, friends, and teachers. *International journal of environmental research and public health, 17*(1), 45.

Kaya, M., & Erdem, C. (2021). Students' well-being and academic achievement: A meta-analysis study. *Child Indicators Research, 14*(5), 1743–1767.

Man, X., Liu, J., & Xue, Z. (2022). Effects of bullying forms on adolescent mental health and protective factors: a global cross-regional research based on 65 countries. *International journal of environmental research and public health, 19*(4), 2374.

Mewafarosh, R., & Agarwal, S. (2022). Influence of demographics variables on students subjective well-being. Journal of Public Affairs, 22, e2749. https://doi.org/10.1002/pa.2749

Ministry of Primary and Secondary Education of the Republic of Indonesia. (2026). Peraturan Menteri Pendidikan Dasar dan Menengah Republik Indonesia Nomor 6 Tahun 2026 tentang Budaya Sekolah Aman dan Nyaman [Regulation of the Minister of Primary and Secondary Education of the Republic of Indonesia Number 6 of 2026 about Safe and Comfortable School Culture]. Ministry of Primary and Secondary Education.

OECD, The Organisation for Economic Co-operation and Development (2022). *PISA 2022 Database: Student questionnaire data file.* https://www.oecd.org/en/data/datasets/pisa-2022-database.html

Phan, H. P., Ngu, B. H., & Alrashidi, O. (2016). Role of Student Well-Being: A Study Using Structural Equation Modeling: A Study Using Structural Equation Modeling. *Psychological Reports, 119*(1), 77–105. https://journals.sagepub.com/doi/10.1177/0033294116656819

Schotanus-Dijkstra, M., Drossaert, C. H., Pieterse, M. E., Boon, B., Walburg, J. A., & Bohlmeijer, E. T. (2017). An early intervention to promote well-being and flourishing and reduce anxiety and depression: A randomized controlled trial. *Internet interventions, 9*, 15–24.

Sharma, R., Singh, P., & Kaushik, A. (2025). Wellbeing Matters: Exploring Levels and Demographic Differences Among Higher Education Students. Current Research *Journal of Social Sciences and Humanities, 8*, 67–75. https://dx.doi.org/10.12944/CRJSSH.8.1.06

World Health Organization. (2025). Corporal Punishment of Children: The public health impact.

# Evaluating the Effectiveness of Gamification Technology in Teaching Science Subjects

Anuarbek Amanov, Khoja Akhmet Yasawi International Kazakh-Turkish University, Kazakhstan
Nurseit Zhunissov, Khoja Akhmet Yasawi International Kazakh-Turkish University, Kazakhstan
Arypzhan Aben, Khoja Akhmet Yasawi International Kazakh-Turkish University, Kazakhstan

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This study is designed to develop an educational platform based on gamification elements to increase the motivation and activity of 7th grade students in teaching natural sciences and to comprehensively evaluate its effectiveness. In the current education system, the decline in interest in natural sciences and the passivity of students are one of the pressing issues. In this regard, during the study, an interactive digital platform based on game mechanics (points, badges, leaderboards, levels, reward system) was developed, which ensures the active participation of students in the learning process. The developed platform was tested in an experimental group among 7th grade students. Within the framework of the research design, a comparative analysis was conducted with traditional teaching methods, and the level of students' academic achievements and interest in the subject was determined. Psychometric tools widely used at the international level were used for the assessment, in particular, the Learning Motivation and Learning Strategies Assessment Questionnaire (MSLQ) and the Intrinsic Motivation Inventory (IMI). In addition, preliminary and final testing methods were used to determine the learning outcomes. The results of the study showed that the use of a gamified platform significantly increased students' cognitive interest in natural sciences, learning activity, and academic performance. Among students, there was a development of self-learning, competitive, and cooperative skills. The proposed gamification approach can be considered an effective tool for revitalizing the educational process and improving the quality of teaching. The results of the study prove the importance of introducing innovative pedagogical technologies in the context of digitalization of education.

*Keywords:* gamification, science subjects, teaching technology, learning motivation, digital education, self-determination theory, academic achievement

iafor
The International Academic Forum
www.iafor.org

## Introduction

The digital transformation in education today has transformed the educational environment by increasing access to interactive content, instant feedback, and improved learning analytics. It has transformed the educational landscape by expanding access to student-centered learning. However, the availability of technology alone does not guarantee active learning. A persistent challenge in secondary school science education (computer science, physics, chemistry) is that students may perceive scientific concepts as abstract, difficult, or disconnected from everyday experience. When instruction relies heavily on explaining and reproducing information, students may become passive, which can weaken interest, persistence, and academic engagement.

Gamification has emerged as one approach for addressing this problem. Deterding et al. (2011) define gamification as the use of game-design elements in non-game contexts. In education, this means retaining curricular goals while incorporating mechanisms such as points, levels, badges, challenges, progress indicators, leaderboards, and rewards. These mechanisms can make progress visible, provide frequent feedback, and create structured goals that encourage learners to continue participating. Importantly, gamification differs from game-based learning: the curriculum does not have to become a complete game; instead, selected game elements are layered onto instructional activities.

The evidence base is broadly encouraging but not uniformly positive. Hamari et al. (2014) concluded that gamification often produces beneficial outcomes, while emphasizing that effects depend strongly on context and users. In a meta-analysis, Sailer and Homner (2020) reported small positive effects on cognitive, motivational, and behavioral outcomes, with the strongest and most stable evidence for cognitive learning. Their analysis also showed that social interaction matters: combinations of competition and collaboration can outperform purely competitive designs. In science education specifically, Kalogiannakis et al. (2021) identified growing use of gamification and emphasized the importance of linking game mechanics to learning objectives rather than treating them as decorative add-ons. A more recent systematic review of primary and secondary education likewise found that research commonly addresses motivation, participation, self-regulation, and engagement, while calling for broader and more rigorous evaluation of school engagement (Ruiz et al., 2024).

This study is conducted with seventh grade students studying science subjects. The purpose of this study is to evaluate a game platform in a digital environment. The platform is designed around a simple motivational loop: students engage in curriculum tasks by playing Kazakh games. They progress through levels and are given opportunities for individual and collaborative achievement. The study compares changes in learning motivation and academic performance between a group using the platform and a control group receiving traditional instruction.

Accordingly, the study addressed three research questions: (1) How does a gamified digital platform affect seventh grade students' motivation and engagement in science? (2) How does the intervention affect academic achievement compared to traditional instruction? (3) What broad learning behaviors, such as self-management, competition, and collaboration, emerge during the gamified intervention?

# Theoretical Background

## Gamification as an Instructional Design Approach

In effective education, gamification is not just about adding rewards to existing tasks. Its value lies in the connection between game mechanics, learning objectives, feedback, and student behavior.

- Points can indicate task completion and effort;
- Badges can indicate mastery or persistence; levels can indicate increasing difficulty;
- Progress bars can support monitoring;
- Leaderboards can introduce social comparison.

When these elements are aligned with curriculum objectives, they can create a clear sense of progress and help students understand what they have done and what they need to learn.

Research also cautions against over-reliance on external reward systems. Leaderboards that repeatedly highlight the same high-performing students can be frustrating to others, and points awarded for superficial actions can distract from conceptual understanding. For these reasons, the platform in this study combined competitive mechanisms with team tasks and multiple forms of recognition. The design aims to make improvement, participation, and mastery visible, rather than just rewarding the highest absolute score.

## Self-Determination Theory and Motivation

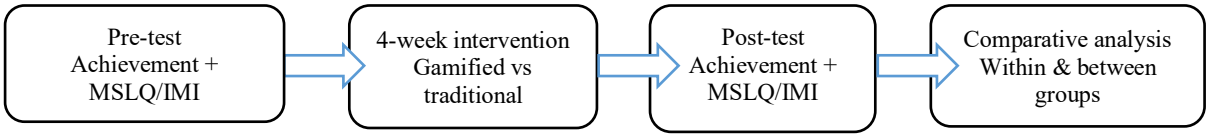
The motivational logic of the platform is consistent with self-determination theory (SDT). Ryan and Deci (2000) explain high-quality motivation through the satisfaction of three basic psychological needs: autonomy, competence, and relatedness. Autonomy concerns the experience of meaningful choice and self-direction; competence involves feeling capable of improving and succeeding; relatedness reflects connection with other people. Gamified learning can support competence through immediate feedback and visible progress, autonomy through choices among tasks or challenge levels, and relatedness through peer recognition and collaborative activities. Sailer et al. (2017) similarly demonstrated that different game-design elements can influence psychological need satisfaction in different ways. This theoretical perspective is useful because it shifts attention from whether students simply “like games” to whether the learning environment supports psychological conditions associated with sustained engagement.

# Materials and Methods

## Research Design and Participants

A quasi-experimental pretest/posttest control group design was used. Participants were seventh-grade students from a science (computer science, physics, chemistry) cohort. The experimental group received instruction via a gamified digital platform, while the control group continued traditional non-gamified instruction. Both groups received the same curriculum over the same period of time. The intervention lasted four weeks. The available conference dataset does not indicate the exact number of participants or group sizes; if required by the conference template, these values should be entered from the participant register before final submission.

**Figure 1**
*Pre-test/Intervention/Post-test Research Design*




## Gamified Digital Platform

The platform incorporated game mechanics directly into science learning activities:

- students earned points for completing assignments, quizzes, and experiments, as well as engaging games;
- badges recognized milestones of mastery and sustained effort;
- levels indicated progression from basic to advanced topics;
- leaderboards made classroom progress visible;
- and rewards unlocked recognition or additional content.

The student dashboard displayed subject progress, recent badges, activity series, and class rankings. The platform therefore transformed otherwise abstract progress information into instant visual feedback that students could track as they engaged.

**Table 1**
*Gamification Elements and Intended Educational Functions*

| Element | Implementation in the platform | Educational function |
|---|---|---|
| Points | Tasks, quizzes, and experiments | Immediate performance feedback and reinforcement |
| Badges | Mastery, effort, and milestones | Recognition of specific achievement and persistence |
| Levels | Progression from basic to advanced units | Visible progression and graduated challenge |
| Leaderboards | Class ranking by accumulated points | Social comparison and healthy competition |
| Rewards | Unlockable recognition/content | Short-term reinforcement linked to achievement |
| Collaborative challenges | Team-based science tasks | Peer support, relatedness, and cooperative problem solving |

## Measurement Instruments

Three forms of measurement were used. First, motivational orientations and learning strategies were assessed using components derived from the Motivated Strategies for Learning Questionnaire (MSLQ). The MSLQ was developed to assess motivation and learning strategies in course contexts and includes dimensions such as goal orientation, task value, self-efficacy, cognitive strategy use, metacognition, and effort regulation (Pintrich et al., 1991). Because the original instrument was designed for college students, any seventh-grade adaptation should be documented carefully, including translation procedures and reliability estimates.

Second, the Intrinsic Motivation Inventory (IMI) was used to capture students' subjective experience of learning activities, particularly interest/enjoyment, perceived competence, effort/importance, and pressure/tension. The IMI has been widely used to operationalize intrinsic motivation and related constructs; its psychometric properties have been examined in prior research (McAuley et al., 1989). In this study, MSLQ- and IMI-derived indicators were summarized into the motivation index reported in the presentation dataset.

Third, identical curriculum-aligned achievement tests were administered before and after the intervention. Using the same assessment framework for both groups enabled direct descriptive comparison of change over time. The combination of achievement measures and motivational instruments allowed the intervention to be evaluated across both cognitive and affective learning outcomes.

### Procedure

The study consisted of four phases. In the initial phase, both groups completed pre-achievement testing and motivation measures. Over the next four weeks, the experimental group was taught the same curriculum using a gaming platform, while the control group was taught traditional teaching. At the end of the intervention, both groups completed the testing structure and motivation measures at the same level. Finally, changes before and after integration were compared within and between groups. Qualitative classroom observations were also used to identify changes in self-management, competitive behavior, and cooperation.

### Statistical Analysis

This conference dataset contains group-level pretest and posttest means, but does not include participant-level scores, sample sizes, standard deviations, or standard errors. Therefore, the analysis presented here is descriptive, not inferential. Our absolute advantage was calculated as a pretest minus a posttest. The descriptive benefit difference was calculated by subtracting the benefit of the control group from the benefit of the experimental group. The relative percent change from baseline was also calculated as (posttest − pretest)/pretest × 100. These calculations do not indicate statistical significance but describe the magnitude of the observed change.

For the final empirical manuscript, participant-level data should be analyzed with methods appropriate to a pre-test/post-test comparison. An independent-samples t-test on change scores can be used when its assumptions are satisfied, but ANCOVA is generally preferable because it compares post-test outcomes while adjusting for baseline differences. Standardized effect sizes such as Cohen's d (with pooled standard deviation) and 95% confidence intervals should accompany p-values. If questionnaire subscales are reported, internal consistency (for example, Cronbach's alpha or omega) should also be provided. These quantities cannot be validly reconstructed from group means alone and were therefore not fabricated in this paper.

## Results

The descriptive data show a consistently larger improvement in the experimental group. For the composite motivation index, the experimental group increased from 58 at pre-test to 84 at post-test, an absolute gain of 26 points. The control group increased from 56 to 63, a gain of 7 points. Thus, the experimental group's descriptive motivation gain exceeded the control

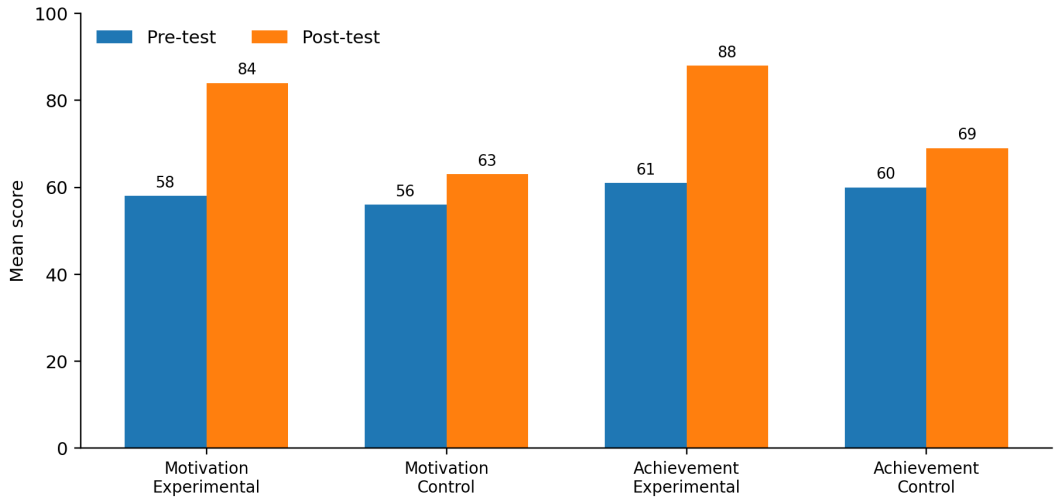
group's gain by 19 points. Relative to baseline, the motivation index increased by 44.8% in the experimental group compared with 12.5% in the control group.

Academic achievement followed the same pattern. The experimental group increased from 61 to 88, a gain of 27 points, whereas the control group increased from 60 to 69, a gain of 9 points. The difference-in-gains was therefore 18 points. Relative to baseline, academic achievement increased by 44.3% in the experimental group compared with 15.0% in the control group. Baseline group means were close for both outcomes, whereas the post-test separation was much larger after the four-week intervention.

**Table 2**
*Descriptive Changes in Motivation and Academic Achievement*

| Outcome | Group | Pre | Post | Gain | Relative change |
|---|---|---|---|---|---|
| Motivation | Experimental | 58 | 84 | +26 | +44.8% |
| Motivation | Control | 56 | 63 | +7 | +12.5% |
| Achievement | Experimental | 61 | 88 | +27 | +44.3% |
| Achievement | Control | 60 | 69 | +9 | +15.0% |

**Figure 2**
*Pre-test and Post-test Mean Scores by Group and Outcome*




Beyond the numerical scores, classroom observations in the experimental group indicated three notable behavioral shifts. First, students showed greater self-directed learning by taking more initiative in pacing and approaching tasks. Second, the leaderboard and progress mechanisms promoted healthy competition focused on improvement and visible progress. Third, team-based challenges supported peer assistance and cooperative problem solving. These qualitative observations are consistent with the platform's design, which combined individual progression with social interaction.

## Discussion

The purpose of this study was to evaluate whether a gamified digital platform could strengthen motivation and achievement in seventh-grade natural science. The descriptive pattern is clear: both groups improved over four weeks, but the gains in the experimental group were substantially larger for both motivation and academic performance. The fact that the control

group also improved is expected because it continued to receive science instruction. The more informative comparison is the difference in the size of the gains, which favored the experimental group by 19 points for motivation and 18 points for achievement.

This pattern is compatible with the broader research literature. Sailer and Homner (2020) found positive average effects of gamification on cognitive, motivational, and behavioral learning outcomes, although the motivational and behavioral effects were less stable in high-rigor subsets. This caution is important for interpreting the present results. The four-week study suggests a promising effect, but the current conference dataset does not provide the participant-level information needed to establish whether the observed differences are statistically reliable or to quantify standardized effect sizes. Accordingly, the findings should be described as substantial descriptive improvements rather than as confirmed statistically significant effects.

The results are also plausible from an SDT perspective. Points, levels, badges, and progress indicators may strengthen perceived competence by making improvement visible. Allowing students to monitor progress and engage with structured challenges may support autonomy. Leaderboards and team challenges create a social dimension that can support relatedness when competition is balanced with cooperation. This balance matters because meta-analytic evidence suggests that competitive-collaborative designs can be more effective than competition alone for motivational and behavioral outcomes (Sailer & Homner, 2020).

For science education, the platform has additional pedagogical relevance. Scientific learning often involves sequential conceptual development, problem solving, testing ideas, and receiving feedback. These characteristics map naturally onto levels, challenges, quests, and progress systems. A well-designed platform can therefore make the logic of scientific progression visible rather than merely adding entertainment. The systematic review by Kalogiannakis et al. (2021) similarly emphasizes the potential of gamification in science education while noting the need for careful design and evaluation.

The observed development of self-directed and cooperative behaviors also extends the interpretation beyond test scores. Recent research on school engagement emphasizes that engagement includes behavioral, emotional, cognitive, and self-regulatory components rather than motivation alone (Ruiz et al., 2024). From this perspective, a useful gamification system should not simply maximize point collection. It should help students set goals, monitor progress, persist with challenging tasks, collaborate productively, and connect feedback to learning decisions.

Several limitations should be acknowledged. The intervention was relatively short, so a novelty effect cannot be excluded. The available dataset does not provide sample size, dispersion measures, or participant-level scores, preventing valid t-tests, ANCOVA, confidence intervals, and Cohen's d calculations. In addition, the MSLQ was originally designed for older learners; the specific seventh-grade adaptation and its psychometric reliability should be reported in a full journal version. Future work should therefore use a larger documented sample, report validated scale properties, extend the intervention period, analyze retention, and investigate whether effects differ according to initial achievement, gender, or learner profile.

## Conclusion

The study demonstrates the practical potential of a gamification-based digital platform for seventh-grade science education. Over a four-week period, the experimental group showed

much larger descriptive improvements than the control group in both the motivation index and academic achievement. Motivation increased by 26 points in the experimental group compared with 7 points in the control group, while achievement increased by 27 points compared with 9 points. The resulting difference-in-gains favored the gamified condition by 19 points for motivation and 18 points for achievement.

These findings suggest that gamification can be most useful when game mechanics are tightly connected to curriculum objectives and when individual progress is combined with cooperative learning. Points, badges, levels, progress tracking, leaderboards, and rewards should function as feedback and self-regulation supports rather than as isolated incentives. The platform's observed influence on self-directed learning, healthy competition, and cooperation further indicates that gamification can contribute to a more active, learner-centered science classroom. A larger study with participant-level statistical analysis is the next step for establishing inferential evidence and standardized effect sizes.

## Acknowledgment

This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant № AP 26193791).

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

Generative artificial intelligence technologies were used only for the translation and grammar checking of this manuscript. In particular, ChatGPT (OpenAI), Claude.ai tools were used to improve the translation and grammar of the text into English, and the readability of the text. B This artificial intelligence tool was not used to generate research ideas, methodology, data analysis, results, or conclusions.

## References

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification." In *Proceedings of the 15th International Academic MindTrek Conference* (pp. 9–15). ACM. https://doi.org/10.1145/2181037.2181040

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In *2014 47th Hawaii International Conference on System Sciences* (pp. 3025–3034). IEEE. https://doi.org/10.1109/HICSS.2014.377

Kalogiannakis, M., Papadakis, S., & Zourmpakis, A.-I. (2021). Gamification in science education: A systematic review of the literature. *Education Sciences, 11*(1), 22. https://doi.org/10.3390/educsci11010022

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport, 60*(1), 48–58. https://doi.org/10.1080/02701367.1989.10607413

Pintrich, P. R., Smith, D. A. F., García, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. National Center for Research to Improve Postsecondary Teaching and Learning, University of Michigan.

Ruiz, J. J. R., Sanchez, A. D. V., & Figueredo, O. R. B. (2024). Impact of gamification on school engagement: A systematic review. *Frontiers in Education, 9*, 1466926. https://doi.org/10.3389/feduc.2024.1466926

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior, 69*, 371–380. https://doi.org/10.1016/j.chb.2016.12.033

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review, 32*(1), 77–112. https://doi.org/10.1007/s10648-019-09498-w

**Contact email:** anuarbek.amanov@ayu.edu.kz

# English Production Learning Cycle Using Generative AI in Image Description Tasks: Focusing on Lowering Psychological Barriers and Promoting Autonomous Noticing

Harumi Kashiwagi, Kobe University, Japan
Min Kang, Kobe University, Japan

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

This study examined how a three-step learning cycle—(1) independent English production, (2) reference to AI-generated samples, and (3) self-correction—affected university students' psychological barriers and "noticing" in AI-generated image description tasks. A total of 129 university students were divided into two groups ($n$ = 65, $n$ = 64) and performed two identical tasks while referring to AI-generated sample descriptions based on two different educational frameworks. Participants rated the AI-generated samples on a 5-point Likert scale across vocabulary, grammar/syntax, and organization, followed by an overall survey. Key findings revealed that: First, over 90% of participants perceived the cycle positively, as the subsequent revision process lowered anxiety about mistakes and fostered psychological security. Second, both groups assigned usefulness ratings of 4.4 or higher on average across all dimensions. While the ratings for vocabulary and grammar/syntax showed no significant differences between the groups, a significant difference was observed in one task, in which structures explicitly describing spatial relationships (e.g., "in the middle," "to the left," or "on the right side") were rated as particularly helpful. Furthermore, over 90% reported a sense of improvement from revising their own work using AI-generated samples as a point of reference rather than merely viewing model answers. Additionally, 60–70% identified AI-generated similar images and texts as effective support resources. These results suggest that sample English sentences generated by generative AI can serve as triggers to promote autonomous awareness in learners, while also providing design principles for psychologically supportive AI-assisted language learning environments.

*Keywords*: generative AI, learning cycle, psychological barriers, noticing, autonomous learning, L2 production

iafor
The International Academic Forum
www.iafor.org

# Introduction

## Research Background

In recent years, driven by growing global demands, higher education English language instruction has increasingly emphasized the development of "production skills"—the ability to proactively express one's own ideas and describe situations—moving beyond the passive reception of knowledge. Specifically, enhancing speaking skills, which directly connect to real-world communication, serves as a central challenge in fostering autonomous language learners.

To effectively enhance these production skills, learners' conscious awareness is crucial for internalizing linguistic information as "intake," rather than merely receiving it passively. In Second Language Acquisition (SLA) research, Schmidt (1990) proposed the "Noticing Hypothesis," positing that subjective "noticing" of linguistic features is a necessary condition for converting input into intake. According to this framework, actively attending to linguistic forms and recognizing the gaps between one's current interlanguage and the target language play a central role in helping learners process subsequent input more attentively and internalize new linguistic knowledge. Regardless of the abundance of input exposure, input is unlikely to be processed as data that promotes language acquisition unless the learner's conscious "noticing" occurs. Therefore, designing an environment that effectively directs learners' attention to appropriate linguistic forms and triggers autonomous "noticing" is a critically important task for modern foreign language instruction, as well as for material development and learning system design.

## Challenges in Foreign Language Production and Speaking

In foreign language production, particularly during speaking tasks that demand fluency and immediacy, learners face several intrinsic challenges at both psychological and cognitive levels.

First, there is the issue of learners' psychological barriers and the accompanying "sense of incompleteness." Psychological factors such as anxiety and lack of confidence can pose significant barriers that hinder both language processing and production. Particularly among English as a Foreign Language (EFL) learners, such as Japanese university students, an excessive fixation on accuracy—and the pervasive fear of making mistakes—raises high psychological hurdles during sentence generation and speaking. Furthermore, in conventional classroom settings, time constraints often force tasks to conclude before learners can adequately reflect on or self-correct their imperfect output. The accumulation of this "sense of incompleteness" diminishes learners' motivation and amplifies their apprehension toward speaking.

Second, there is difficulty in connecting "noticing" to "autonomous self-correction." In speaking instruction, presenting immediate, multidimensional feedback—such as vocabulary, grammar/syntax, and organization—for each individual learner is extremely difficult due to structural classroom constraints. Particularly in tasks that involve expressing structured content, it is crucial to direct learners' attention not merely to isolated vocabulary or grammatical corrections, but also to organizational aspects, such as the logical sequence of information and spatial relationships needed to describe a situation objectively. However, simply providing pre-prepared "model answers" often leads learners to perceive them as the

"one and only correct answer." This approach has clear limitations in promoting autonomous noticing, where learners must actively consider how to refine their own original production.

**Designing a Learning Cycle Utilizing AI-Generated Sample Answers**

To address these psychological and cognitive challenges in speaking outlined above, the educational application of generative AI has garnered significant attention. Generative AI can flexibly produce a wide range of English texts tailored to task contexts and learning goals, making it a powerful tool for supporting autonomous language learning. In language teaching practice, efforts to provide opportunities for language production (in both speaking and writing) are being actively implemented through AI technologies. These include AI-assisted speaking practice using conversational AI tools (Shen et al., 2025) and AI-supported writing instruction (Chen, 2026).

However, researchers have pointed out that relying on AI for direct error correction or immediate provision of correct answers poses a notable challenge: learners tend to accept AI output uncritically, which hinders active revision and deep "noticing." Although these insights stem primarily from second language (L2) writing, the process of monitoring one's language production and underlying knowledge through written form—and subsequently engaging in revision—plays a crucial cognitive role in building the foundation for future speaking proficiency (oral production). Specifically, Chen (2026) highlights that while generative AI can enhance planning and evaluative strategies, easy access to AI validation often leads to a "monitoring displacement" mechanism—where learners outsource real-time self-regulation during drafting, consequently weakening their motivation for internal monitoring. Furthermore, Chen and Liu (2026) caution that when learners use AI as an outcome-oriented tool, they tend to outsource their cognitive evaluation and agency to the system. Over time, this passive reliance diminishes their sense of ownership and can erode their capacity for self-regulation and strategic planning. Consequently, when provided with direct corrections or model answers, learners tend to accept or copy outputs uncritically rather than actively engaging with the feedback.

To address these challenges, Zare et al. (2025) emphasize that providing long-term, scaffolded support to help learners become independent producers—rather than allowing them to rely passively on generative AI—is essential for sustaining motivation and self-regulation. As a concrete approach to task design, Pust (2026) demonstrates that an inductive framework—in which students actively manipulate digital tools to discover language rules and undergo iterative text revision—effectively fosters learner engagement. Furthermore, Michta and Frankenberg-Garcia (2026) demonstrate that, unlike systems that automatically rewrite text, a design that allows users to select suggestions while retaining their agency enables learners to make autonomous choices—such as rejecting unnecessary edits—thereby fostering noticing and vocabulary learning.

Building upon these insights, this study designed a learning cycle using an "image description task" as a concrete means of mitigating the psychological and cognitive challenges learners face in speaking. In this cycle, immediately after learners produce their own English sentence, they are presented with pre-generated "sample answers" by generative AI based on two distinct educational frameworks. These samples serve as a reference for self-correction and re-production.

This approach intentionally adopts a text-based (writing) framework for sentence production, aiming to enhance the quality of "noticing"—the foundation for subsequent speaking—by allowing learners to visually and carefully compare their own English with the AI sample answer. Crucially, the sample answers provided in this cycle are neither passive feedback through direct AI error correction nor fixed, conventional model answers that learners often perceive as the "one and only correct answer." This design acknowledges that AI-generated text is not always optimal, thereby encouraging learners to critically evaluate the provided options rather than passively adopt them.

Positioning controlled AI-generated text as merely "one example" and encouraging learners to reference it autonomously—as a guide to incorporate more appropriate expressions and structures while using their own English as a baseline—is expected to elicit autonomous noticing not only of vocabulary and grammar/syntax, but also of discourse structure and an active revision process. We consider this process of contrasting one's own English with a sample that may contain imperfections, rather than receiving individualized direct correction, to hold significant educational value for autonomous learning. Furthermore, establishing a framework premised on revision rather than expecting flawless output from the start is anticipated to alleviate learners' sense of incompleteness, mitigate anxiety over failure, and foster psychological safety. Task design that accounts for such psychological safety and autonomy may lead learners to positively embrace generative AI as a partner in self-directed learning and cultivate a sustained attitude toward utilizing AI.

## Research Objectives

Based on the issues outlined above, the purpose of this study is to implement and evaluate a "sentence production learning cycle" among Japanese university students to foster future speaking proficiency. This cycle incorporates self-correction through autonomous reference to pre-generated AI sample answers in an image description task. Specifically, this study addresses the following two Research Questions (RQs):

RQ1: How does a learning cycle premised on revision, in which learners autonomously reference and compare AI-generated "sample answers (as one of many possibilities)," affect the alleviation of learners' sense of incompleteness and the cultivation of psychological safety toward sentence production?

RQ2: How do learners perceive the utility of AI sample answers across linguistic and structural dimensions (vocabulary, grammar/syntax, and organization), and how does this perception contribute to their linguistic noticing and expectations for further generative AI-assisted learning?

Connecting linguistic utility with broader learning outcomes in RQ2 is essential because evaluating the perceived utility across specific linguistic dimensions serves as the immediate cognitive trigger that drives learners' linguistic noticing and shapes their subsequent expectations for AI-assisted learning. Through these investigations, this study seeks to explore the potential of leveraging generative AI to support autonomous learning and offer concrete implications for building a foundation for future speaking proficiency.

# Methods

## Participants

The participants in this study were 129 first-year Japanese university students. Based on the specific types of AI-generated sample sentences provided, they were assigned to two independent groups under different experimental conditions: Group A ($n$ = 65) and Group B ($n$ = 64).

## Tasks and Procedures

In this study, an "English picture description task," which required participants to describe AI-generated illustrations in English, was conducted across two separate sessions. The illustrations depicted scenes of a "park" and a "supermarket," respectively. Both the task illustrations and the AI-generated sample sentences were generated using Gemini 3.6 Flash. To address the purpose of this study, the tasks were carried out according to the three-step "sentence production learning cycle" outlined below:

1. **Independent English Production**: After viewing the presented illustration, participants first produced (wrote) their initial descriptive sentences entirely on their own.
2. **Reference to Sample Sentences**: Next, pre-controlled sample sentences generated by Gemini 3.6 Flash were presented, and participants referred to and compared them with their own sentences. Through specific prompt instructions, these sample sentences were developed based on two different educational frameworks: the CEFR A2 level and the first-year high school level. Specifically, one framework was based on the Common European Framework of Reference for Languages (CEFR) A2 level, an international language proficiency standard, while the other was aligned with the "first-year high school level" from the domestic national curriculum guidelines. The assignment for each group and session is presented in Table 1.
3. **Self-Revision and Reproduction**: Based on the insights gained from referencing the sample sentences, participants revised their initial sentences and produced the final version of their English text.

**Table 1**
*The Assignment of AI-Generated Sample Sentences for Each Group and Session*

| Session (Topic) | Group A ($n$ = 65) | Group B ($n$ = 64) |
|---|---|---|
| Session 1 (Park) | 1st-year High School level | CEFR A2 level |
| Session 2 (Supermarket) | CEFR A2 level | 1st-year High School level |

*Note.* $n$ = number of participants.

Within this learning cycle, the sentences provided by the generative AI were not positioned as a "model answer" that learners were expected to imitate word for word. Instead, they were treated as one possible "sample answer" generated by AI for the same illustration. While using their initial draft as the primary foundation, participants were instructed to actively and autonomously reference this sample text as a comparative resource to incorporate more appropriate expressions and structures.

## Data Collection and Analysis

To investigate the research questions, two distinct types of questionnaire surveys were administered: one immediately after each task, and a post-activity questionnaire after completing both sessions. First, immediately following each description task, participants evaluated the usefulness of the presented sample answer. Specifically, they rated the extent to which the sample answer served as a helpful reference for their self-revision across three distinct domains—lexical, grammatical/syntactic, and organizational—using a 5-point Likert scale (see Table 2).

Furthermore, after completing both learning sessions, a post-activity questionnaire was administered to comprehensively measure the overall influence of this learning cycle on the learners. This questionnaire consisted of two primary domains—psychological/attitudinal aspects and learning/linguistic aspects—and similarly utilized a 5-point Likert scale to assess participants' perceptions (see Table 3).

For the statistical analysis, the Wilcoxon rank-sum test was used to investigate whether statistically significant differences existed between the ratings of the two groups.

**Table 2**
*Questionnaire Items Administered Immediately After Each Task*

| **Usefulness of Sample Sentences for Self-Revision** |
|---|
| Q1. Domain: Lexical<br>When revising your own English sentences, did you find the use of words and vocabulary expressions in the sample answer helpful?<br>1. Not helpful 2. Not very helpful 3. Neutral 4. Somewhat helpful 5. Helpful |
| Q2. Domain: Grammatical/Syntactic<br>When revising your own English sentences, did you find the use of grammar and syntactic structures in the sample answer helpful?<br>1. Not helpful 2. Not very helpful 3. Neutral 4. Somewhat helpful 5. Helpful |
| Q3. Domain: Organizational<br>When revising your own English sentences, did you find the overall structure and text organization of the sample answer helpful?<br>1. Not helpful 2. Not very helpful 3. Neutral 4. Somewhat helpful 5. Helpful |

**Table 3**
*Post-activity Questionnaire Items*

| **Learners' Perceptions of the Sentence Production Learning Cycle** |
|---|
| **Domain: Psychological /Attitudinal** |
| Q1. The method of revising and resubmitting my initial drafts based on the sample answers was preferable to simply creating and submitting them. |
| 1. Disagree 2. Somewhat disagree 3. Neither agree nor disagree 4. Somewhat agree 5. Agree |
| Q2. It was beneficial that the process of revision and resubmission prevented the learning task from ending with the submission of incomplete sentences. |
| 1. Disagree 2. Somewhat disagree 3. Neither agree nor disagree 4. Somewhat agree 5. Agree |
| **Domain: Learning /Linguistic** |
| Q3. Revising my drafts based on the sample answers helped me notice vocabulary and expressions that I could not initially produce on my own. |
| 1. Disagree 2. Somewhat disagree 3. Neither agree nor disagree 4. Somewhat agree 5. Agree |
| Q4. Using my own drafts as a foundation for revision, rather than merely reading the sample answers, will contribute to the improvement of my English skills. |
| 1. Disagree 2. Somewhat disagree 3. Neither agree nor disagree 4. Somewhat agree 5. Agree |
| Q5. It would have been better if similar illustrations and sample answers explaining them had been provided before I created my initial drafts. |
| 1. Disagree 2. Somewhat disagree 3. Neither agree nor disagree 4. Somewhat agree 5. Agree |

## Results

### Effects on Psychological Aspects (RQ1)

Regarding RQ1, the post-activity questionnaire results (Table 4) revealed that participants in both groups responded favorably to the proposed learning cycle. Specifically, over 90% of the learners in both groups evaluated the revision-and-resubmission cycle positively.

For Q1, which evaluated the method of revising and resubmitting initial drafts based on sample answers rather than simply submitting them, the vast majority of participants in both groups agreed. Specifically, 25% of the participants in Group A selected "Somewhat agree," and 69% selected "Agree," totaling 94%. In Group B, 52% selected "Somewhat agree," and 48% selected "Agree," representing a combined total of 100% (Group A: $M = 4.6$, $SD = 0.8$; Group B: $M = 4.5$, $SD = 0.5$).

Similarly, for Q2, which addressed the benefit of preventing the learning task from ending with the submission of incomplete sentences, a substantial majority in both groups viewed the cycle favorably. Specifically, 25% of the participants in Group A selected "Somewhat agree," and 74% selected "Agree," totaling 99%. In Group B, 36% selected "Somewhat agree," and 64% selected "Agree," representing a combined total of 100% (Group A: $M = 4.7$, $SD = 0.6$; Group B: $M = 4.6$, $SD = 0.5$).

**Table 4**
*Participants' Responses to Post-activity Questionnaire Items*

| Group A | 1. Disagree | 2. Somewhat disagree | 3. Neither agree nor disagree | 4. Somewhat agree | 5. Agree |
|---|---|---|---|---|---|
| Q1 | 1.5% | 3% | 1.5% | 25% | 69% |
| Q2 | 1% | 0% | 0% | 25% | 74% |
| Q3 | 2% | 0% | 0% | 18% | 80% |
| Q4 | 2% | 0% | 3% | 37% | 58% |
| Q5 | 0% | 20% | 8% | 38% | 34% |

| Group B | 1. Disagree | 2. Somewhat disagree | 3. Neither agree nor disagree | 4. Somewhat agree | 5. Agree |
|---|---|---|---|---|---|
| Q1 | 0% | 0% | 0% | 52% | 48% |
| Q2 | 0% | 0% | 0% | 36% | 64% |
| Q3 | 0% | 0% | 0% | 19% | 81% |
| Q4 | 0% | 0% | 8% | 42% | 50% |
| Q5 | 0% | 16% | 20% | 34% | 30% |

**Linguistic and Organizational Utility (RQ2)**

Regarding RQ2, data obtained from both the questionnaire items administered immediately after each task and the post-activity questionnaire demonstrated that the participants deeply recognized the linguistic and organizational benefits of the proposed learning cycle.

Regarding the utility of the sample answers, the immediate questionnaire results indicated a high evaluation across all three targeted aspects—lexical (vocabulary), grammatical/syntactic, and organizational (discourse structure). The mean scores exceeded 4.4 on a 5-point Likert scale across both groups and tasks, ranging from 4.6 to 4.8 in Group A and from 4.4 to 4.7 in Group B (Table 5). While no statistically significant differences were found between the two groups regarding the lexical and grammatical/syntactic aspects at the 0.05 significance level, a significant difference between Groups A and B emerged in the organizational aspect for the specific "park" task (Wilcoxon rank-sum test, $p = .019$).

**Table 5**
*Participants' Responses to the Questionnaires Administered Immediately After Each Task*

| Park | | 1. Not helpful | 2. Not very helpful | 3. Neutral | 4. Somewhat helpful | 5. Helpful |
|---|---|---|---|---|---|---|
| lexical | Group A | 3% | 0% | 0% | 12% | 85% |
| | Group B | 5% | 0% | 2% | 23% | 70% |
| grammatical/ syntactic | Group A | 5% | 0% | 2% | 18% | 75% |
| | Group B | 4.7% | 0% | 7.8% | 25% | 62.5% |
| organizational | Group A | 3% | 0% | 0% | 15% | 82% |
| | Group B | 6% | 0% | 8% | 22% | 64% |

| Supermarket | | 1. Not helpful | 2. Not very helpful | 3. Neutral | 4. Somewhat helpful | 5. Helpful |
|---|---|---|---|---|---|---|
| lexical | Group A | 0% | 0% | 0% | 18% | 82% |
| | Group B | 0% | 2% | 0% | 28% | 70% |
| grammatical/ syntactic | Group A | 0% | 0% | 2% | 29% | 69% |
| | Group B | 0% | 1.6% | 4.7% | 31.2% | 62.5% |
| organizational | Group A | 0% | 0% | 3% | 31% | 66% |
| | Group B | 0% | 1.6% | 1.6% | 20% | 76.8% |

A comparative analysis of the two sample English texts describing the park scene (Table 6) identified clear structural differences between the conditions. The Group A sample utilized a systematic spatial progression—guiding the reader's focus seamlessly from the overall scene to the middle, left, and right areas using explicit locative markers. In contrast, while the Group B sample also adopted a standard top-down organization and attempted spatial framing using "In the middle," its spatial signposting remained limited. Rather than guiding the reader across the entire scene, it concentrated various elements—such as the playground, path, benches, trash can, and flowers—within a single central zone. Consequently, the description of this central area relied primarily on listing specific amenities through repetitive structures (e.g., "There is/are..."), ultimately resulting in a feature-centric listing rather than a systematic spatial progression. Specifically, in Group A, the inclusion of explicit spatial markers—such as "in the middle," "to the left," or "on the right side"—was found to be highly instrumental in guiding the learners' paragraph organization. This suggests that textual structures explicitly highlighting spatial arrangements are perceived as more informative and useful by learners when organizing their descriptive writing. Explaining these structural differences is pedagogically meaningful because distinct framework designs can direct learners' attention to specific linguistic features during self-correction before examining their overall impact on learners.

Furthermore, the post-activity questionnaire provided deeper insights into the learners' awareness of their improvement and the task design. For Q3 in Table 4, which examined whether revising their drafts based on the sample answers helped learners notice vocabulary and expressions they could not initially produce on their own, a substantial majority in both groups responded positively. Specifically, 18% of the participants in Group A selected "Somewhat agree," and 80% selected "Agree," totaling 98%. In Group B, 19% selected "Somewhat agree," and 81% selected "Agree," representing a combined total of 100% (Group A: $M = 4.8$, $SD = 0.6$; Group B: $M = 4.8$, $SD = 0.4$). Similarly, for Q4 (Table 4), which evaluated the effect of using their own sentences as a foundation for revision, rather than merely reading the sample answers, the participants reported strong perceived benefits, recognizing it as an effective way to improve long-term English skills. Specifically, 37% of the participants in Group A selected "Somewhat agree," and 58% selected "Agree," totaling 95%. In Group B, 42% selected "Somewhat agree," and 50% selected "Agree," representing a combined total of 92% (Group A: $M = 4.5$, $SD = 0.7$; Group B: $M = 4.4$, $SD = 0.6$).

Interestingly, the responses to Q5 (Table 4) revealed a clear learner preference regarding the task sequence. Approximately 60% to 70% of the participants in both groups (72% in Group A and 64% in Group B) selected either "Somewhat agree" or "Agree," indicating a distinct preference for having similar illustrations accompanied by sample answers provided as

references prior to creating their initial drafts (Group A: $M = 3.9$, $SD = 1.1$; Group B: $M = 3.8$, $SD = 1.0$).

**Table 6**
*Comparison of AI-Generated Sample Texts for the Park Scene Task*

| Group A: 1st-Year High School Level | Group B: CEFR A2 Level |
|---|---|
| This is a beautiful and sunny park scene. The image shows a playground and a nice, green area with lots of trees. | This is a picture of a nice park on a sunny day. |
| **In the middle**, there is a winding dirt path. On the path's edge, you can see a wooden bench, and another bench is further back. | I can see a lot of green trees and grass. The sky is light blue and there are a few white clouds. |
| **To the left**, there's a big tree and a trash can. | **In the middle**, there is a playground. It has a slide that is orange and a swing set that is blue. There is a brown path that goes through the grass. I also see two wooden benches where people can sit and a black trash can. |
| **On the right side**, there is a play structure with an orange slide and a set of swings. The sky is blue with some fluffy white clouds. | There are some yellow flowers near the front bench. |
| Overall, the picture is colorful and peaceful, like a perfect day outside. | It looks like a good place to play and relax. |

## Discussion

### Psychological Comfort and Safety (RQ1)

Regarding RQ1, the post-activity questionnaire results suggested that over 90% of the participants evaluated the proposed learning cycle very positively. This highly positive self-reported response may not be solely attributable to the novelty or surface-level appeal of using generative AI. Instead, it implies a potential underlying psychological factor associated with the specific task design.

Furthermore, the explicit structure of repeated revision may have helped mitigate learners' anxiety regarding failure. In conventional production tasks, learners often experience a sense of incompleteness or frustration when they cannot express their ideas perfectly on their initial attempt. However, by positioning the sample answers not as an absolute standard of correctness but as a supportive resource for secondary revision, the current learning cycle may have helped reduce this psychological burden. This design could have fostered a sense of psychological safety, potentially transforming hesitation into a constructive noticing experience and enabling learners to engage with the tasks with less apprehension about making mistakes.

### Textual Scaffolding and Learners' Insights (RQ2)

Regarding RQ2, the quantitative analysis of the perceived usefulness of the AI-generated sample answers indicated consistently high ratings. Mean scores exceeded 4.4 on a 5-point Likert scale across all criteria ("vocabulary," "grammar/syntax," and "organization") for both

groups and tasks. However, while no statistically significant differences were observed between the groups in terms of "vocabulary" and "grammar/syntax," a statistically significant difference was found solely in "organization" within a specific task (the "park" task).

A qualitative inspection of the specific text characteristics provided in that task offers a plausible explanation for this variance. The target task involved a descriptive scene of a park, where providing precise spatial descriptions could render the explanation clearer and more coherent. Notably, the sample provided to Group A for the park scene prominently incorporated explicit spatial markers—such as "in the middle," "to the left," and "on the right side"—guiding the reader through a systematic spatial progression. In contrast, while the sample for Group B attempted spatial framing using "In the middle," its spatial signposting remained limited, concentrating various elements within a single zone and ultimately resulting in a more feature-centric listing. The explicit presentation of spatial arrangements in Group A may have functioned as a more effective textual scaffolding. Consequently, this comprehensive visual-spatial support likely provided learners with a clearer structural framework, which may explain why a statistically significant difference was observed specifically in their evaluation of discourse organization.

Furthermore, the survey findings provide valuable insights for future task designs incorporating generative AI. Over 90% of the participants reported that they did not treat the AI samples as mere "model answers" to copy blindly. Instead, they used them as a reference to refine their original writing, which may have enhanced their sense of learning ownership and self-improvement. Additionally, approximately 60% to 70% of the participants in both groups expressed that having access to similar illustrations and corresponding AI-generated sample texts as a pre-task reference—which was not provided in the current study—would have provided useful support. This preference for pre-task materials suggests that integrating complementary visual and textual resources beforehand could potentially enhance the scaffolding effects of AI-assisted language learning environments.

## Conclusion

### Summary of Findings

This study investigated the potential psychological and pedagogical scaffolding effects of integrating AI-generated resources, specifically descriptive illustrations and corresponding sample texts, into university-level English production tasks. Regarding RQ1, the findings suggest that the pedagogical design incorporating repeated revision appears to have helped mitigate learners' anxiety regarding failure. By positioning the AI samples not as an absolute model but as a supportive resource for secondary revision, this learning cycle likely fostered a sense of psychological safety. Consequently, it may have transformed potential anxiety into a constructive noticing experience, thereby enabling learners to engage with the production tasks with reduced apprehension about making mistakes.

For RQ2, the quantitative evaluation revealed that both groups perceived the AI samples as highly useful across all domains—including vocabulary, grammar, and organization—yielding mean scores exceeding 4.4 out of 5 in both tasks. Notably, while no significant differences were observed between the groups in vocabulary and grammar, a statistically significant difference emerged in the domain of organization within a specific descriptive task. Textual analysis suggested that the inclusion of explicit spatial markers (e.g., "in the

middle," "to the left," or "on the right side") in the AI texts may have served as effective textual scaffolding, which may explain this specific variance.

Furthermore, the survey findings suggested a high degree of learner autonomy, as over 90% of the participants reported that they did not treat the AI samples as a mere template for blind copying. Instead, they perceived them as a reference to refine their original writing, potentially enhancing their sense of learning ownership.

## Pedagogical Implications

The findings of this study offer important pedagogical insights for integrating generative AI into foreign language education, particularly regarding the instructional design of production tasks. First, the study suggests the potential benefits of shifting the instructional focus from product-oriented AI evaluation to process-oriented psychological scaffolding. Rather than utilizing AI primarily as an automated grader or a source of instant correction, educators might consider adopting a structured revision cycle that incorporates AI-generated samples as supportive resources. By framing AI-generated texts as a flexible reference rather than an absolute model, this approach may help lower learners' affective filters. Consequently, such a design could foster a supportive environment in which students feel safer to engage with the target language, potentially transforming errors into opportunities for constructive noticing and linguistic growth.

Second, the results suggest that the efficacy of AI scaffolding may be closely linked to the specific structural and linguistic design of the AI-generated sample texts. Rather than merely presenting generic AI-generated models, educators could deliberately construct or prompt AI samples that contain explicit text-structural and discourse-level cues (e.g., spatial markers) to guide learners. Furthermore, an instructional design that encourages students to reflect on and refine their own production based on these AI references may further promote learner autonomy. When structured in this manner, tasks have the potential to enable AI samples to serve as a supportive scaffold for self-directed language learning.

## Limitations and Recommendations for Future Research

Although these findings offer meaningful insights into AI-supported production tasks, several limitations should be acknowledged. First, the participants in this study were limited to Japanese university freshmen from a single institution, representing a relatively narrow range of proficiency. Therefore, the generalizability of these findings to learners with different cultural backgrounds, age groups, or proficiency levels remains to be further explored. Future research incorporating a more diverse sample of participants would be beneficial to examine the extent to which the proposed task design can be applied across varied educational contexts.

Second, this study focused primarily on immediate psychological and behavioral responses within a set of structured tasks. Longitudinal investigations are needed to clarify whether the observed reductions in affective filters and proactive revision behaviors can be sustained over an extended period.

Third, the evaluation in this study relied primarily on learners' subjective perceptions and self-reported data. While these insights are invaluable for understanding affective variables, future research could benefit from incorporating objective measures—such as evaluating

actual pre- and post-task linguistic accuracy, syntactic complexity, and lexical range—to investigate the potential impact of this AI-scaffolded approach on learners' actual language development.

Finally, while the quantitative findings suggested that students tended to actively revise their drafts rather than blindly copy the AI references, qualitative investigations—such as analyzing process log data or conducting retrospective interviews—could provide more nuanced insights into the underlying cognitive processes through which learners interact with and internalize the AI-generated structural and linguistic cues.

## Acknowledgments

This work was supported by JSPS Grant-in-Aid for Scientific Research (C) (Grant No. JP25K06595).

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the authors used Gemini 3.6 Flash to improve language and readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication. Additionally, generative AI technology (Gemini 3.6 Flash) was utilized as part of the research methodology to generate the sample English texts and images analyzed in this study; specific details are described in the Tasks and Procedures subsection of the Methods section.

## References

Chen, M.-R. A. (2026). Generative AI, visual scaffolds, and metacognition in EFL writing. *Language Learning & Technology*, *30*(1), 1–25. https://doi.org/10.64152/10125/73695

Chen, P.-H. P., & Liu, Y. (2026). Distributed agency in AI-assisted L2 writing. Language Learning & Technology, *30*(1), 1–20. https://doi.org/10.64152/10125/73674

Michta, T., & Frankenberg-Garcia, A. (2026). The flow must go on: Evaluating the impact of a collocation writing assistant on self-revision. *ReCALL*, *38*(1), 57–74. https://doi.org/10.1017/S0958344025100293

Pust, D. (2026). Learning grammar the explorative way: Integrating interactive grammar animations into CALL. *Language Learning & Technology*, *30*(1), 1–21. https://doi.org/10.64152/10125/73667

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129

Shen, J., Qian, B., Ma, Q., & Deris, F. D. (2025). Exploring the impact of AI-assisted speaking practice on EFL learners' speaking proficiency and motivation: A mixed-method study. *International Journal of Academic Research in Progressive Education and Development*, *14*(4), 43–52. http://dx.doi.org/10.6007/IJARPED/v14-i4/26672

Zare, J., Al-Issa, A., & Ranjbaran Madiseh, F. (2025). Interacting with ChatGPT in essay writing: A study of L2 learners' task motivation. *ReCALL*, *37*(3), 385–402. https://doi.org/10.1017/S0958344025000035

**Contact email:** kasiwagi@kobe-u.ac.jp

# Exploring the Use of a Culturally Responsive Pedagogy in Practice: Perception and Knowledge Performance With the CARE² Model

Karen K. Fujii, Niels Brock Copenhagen Business College, Denmark
Natalie Perez, University of Hawaii, United States

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

International learning environments are a defining feature of contemporary higher education and require instructional approaches that intentionally address cultural diversity to support student success (Leask, 2015). This study examines the implementation of the Fujii and Perez (2025) CARE² model, a culturally responsive pedagogical framework, in undergraduate marketing courses at a higher education institution in Denmark. The study draws on multiple data sources, including learner surveys on CARE² model perceptions, self-reported technology capability measures collected pre-post, knowledge-based assessments, and final exam performance. The findings indicate strong learner perceptions of the CARE² model ($M = 4.07/5$, $SD = 0.60$), with Instructor Support and Meaningful Content rated highest, and Cultural Relevance and Timely Feedback comparatively lower. Significant improvements were observed in 8 of 12 technology skill areas, particularly in submitting online assessments, creating presentations, and using presentation tools ($p < .05$ to $p < .001$), alongside a substantial reduction in technology-related anxiety. Academic performance averaged 55.7% ($SD = 12.0$), with a 67.6% pass rate. Pre-intervention technology skills were not associated with academic performance. Overall, findings indicate that the CARE² model is associated with positive learner perceptions, meaningful growth in self-reported technological capabilities, and improved confidence in navigating digital learning environments, supported by multiple forms of assessment data.

*Keywords:* culturally-responsible pedagogy, learner experience, educational evaluation, CARE² model, international learning environments

iafor
The International Academic Forum
www.iafor.org

## Introduction

The internationalization of higher education has created increasingly diverse learning environments characterized by linguistic, cultural, and technological differences (Leask, 2015). While this diversity enriches higher education, it also presents pedagogical challenges, particularly for learners with varying levels of experience with digital technologies, Western assessment practices, and the academic conventions of host institutions. Consequently, educators must design learning experiences that are both academically rigorous and responsive to learners' diverse cultural and educational backgrounds.

Culturally responsive pedagogy (CRP) offers a well-established framework for addressing these challenges by affirming learners' cultural identities, drawing on prior knowledge, and supporting their participation in new academic communities (Gay, 2000; Ladson-Billings, 1995). Although CRP has gained increasing attention in higher education, empirical evaluations remain limited, particularly in undergraduate settings and non-English-speaking host countries. Existing research has largely been theoretical or descriptive, with few studies examining the effects of culturally responsive interventions on multiple outcomes, including technological capability, academic performance, and learner confidence.

The CARE² model (Fujii & Perez, 2025) was developed to operationalize culturally responsive pedagogy in higher education. Derived through iterative qualitative analysis across multiple learner cohorts, the framework organizes eight instructional principles into four dimensions: Culture and Collaboration, Agency and Agility, Relevance and Repetition, and Engagement and Evaluation. Integrating culturally responsive pedagogy with cognitive science, learner engagement, and formative assessment, the model provides a structured framework for designing instruction that supports diverse learners.

This study evaluates the implementation of the CARE² model in undergraduate marketing courses at a higher education institution in Denmark serving primarily learners from Nepal and Bangladesh. Because many learners entered the course with limited experience using digital technologies and laptop-based assessments, the study examines whether the model was associated with improvements in technological capability, academic performance, and learner perceptions. The following research questions guided the study:

1. How do undergraduate learners in an international higher education context perceive the components of the CARE² model, and which dimensions are rated most and least favorably?
2. To what extent are learners' self-reported technological capabilities associated with significant change following a CARE² model-informed instructional intervention?
3. What is the relationship between learners' pre-intervention technology skill levels and their academic performance on knowledge-based assessments?

## Literature Review

### Internationalization and the Challenge of Cultural Diversity in Higher Education

The internationalization of higher education has created increasingly diverse learning environments, presenting both opportunities and pedagogical challenges. Leask (2015) argues that meaningful internationalization requires more than increasing international enrollment; it requires adapting curriculum, pedagogy, and assessment to support equitable participation. Without such changes, institutions risk reinforcing existing inequities. Research shows that

international learners, particularly those from South and Southeast Asia, often face challenges beyond language, including unfamiliarity with student-centered learning, expectations of independent critical thinking, and digitally mediated assessment systems common in Western universities (Montgomery, 2010; Ryan, 2011). For learners from educational contexts emphasizing teacher-directed instruction and paper-based assessment, this transition can be a significant academic and technological adjustment.

## Culturally Responsive Pedagogy (CRP)

Culturally responsive pedagogy (CRP), grounded in the work of Gay (2000) and Ladson-Billings (1995), challenges deficit-based views of learner diversity by recognizing cultural knowledge, prior experiences, and ways of learning as instructional assets. Gay (2000) emphasized the use of learners' cultural frames of reference to make learning more relevant and effective, while Ladson-Billings (1995) highlighted academic success, cultural competence, and critical consciousness as central goals of responsive teaching. Although initially developed within U.S. K–12 contexts, CRP has been extended to higher education and international settings, where scholars emphasize the importance of recognizing learners' cultural, linguistic, and social resources (Hockings, 2010; Mthethwa-Sommers, 2014). Research suggests that culturally responsive approaches can enhance both learner experiences, including motivation and belonging, and academic outcomes, particularly for learners whose backgrounds differ from dominant institutional norms (Gay, 2018; Sleeter, 2012). This dual emphasis on equity and achievement provides the theoretical foundation for the present study.

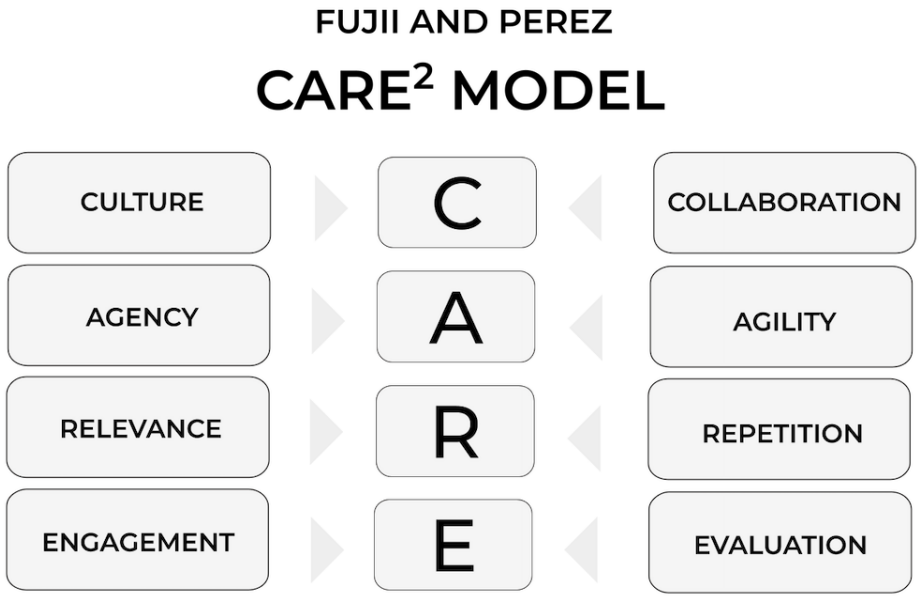
## The CARE² Model: A New Culturally Responsive Pedagogy

The CARE² model, developed by Fujii and Perez (2025), is a theoretically grounded and empirically derived framework for designing culturally responsive learning experiences in higher education. It comprises eight instructional principles organized into four dimensions: Culture and Collaboration, Agency and Agility, Relevance and Repetition, and Engagement and Evaluation (see Figure 1). Together, these dimensions balance cultural affirmation with collaborative learning, learner autonomy with instructional adaptability, meaningful contextualization with cognitive reinforcement, and sustained engagement with formative assessment.

- **Culture and Collaboration (C)** positions learners' cultural backgrounds as instructional assets rather than barriers. Consistent with Gay (2000), the model encourages educators to draw on learners' prior experiences and cultural knowledge while fostering collaborative learning communities that support academic engagement across diverse backgrounds (Johnson & Johnson, 2009).
- **Agency and Agility (A)** emphasizes both learner autonomy and instructional flexibility. Agency promotes learner ownership of learning, enhancing motivation and engagement (Reeve, 2012; Ryan & Deci, 2000), while agility highlights the instructor's ability to adapt instruction in response to learners' evolving needs. Together, these principles reflect culturally responsive teaching as an iterative and adaptive practice.
- **Relevance and Repetition (R)** integrates culturally meaningful instruction with evidence-based learning principles. Connecting content to learners lived experiences enhances motivation (Hammond, 2015; Keller, 1987), while repeated exposure across varied contexts strengthens long-term retention and transfer (Cepeda et al., 2006; Sweller, 1988). This combination is particularly valuable for international learners navigating unfamiliar linguistic, cultural, and technological environments.

- **Engagement and Evaluation (E)** views assessment as an integral component of learning rather than a separate endpoint. The model promotes behavioral, cognitive, and affective engagement (Fredricks et al., 2004) while emphasizing formative assessment and feedback to support learning (Hattie & Timperley, 2007). For international learners, making assessment expectations explicit and dialogic further enhances learning and equity (Carless, 2006).

**Figure 1**
*Fujii & Perez CARE² Model*




**Digital Equity and Technological Capability in International Higher Education**

Digital equity has become an increasingly important consideration in international higher education, reflecting disparities in learners' access to, familiarity with, and confidence using digital technologies. The rapid expansion of digitally mediated learning, accelerated by the COVID-19 pandemic, has highlighted how differences in technological preparedness can disadvantage learners with limited prior experience using computers, learning management systems, and digital productivity tools (Bond et al., 2021; Czerniewicz et al., 2020). These disparities are shaped less by learners' academic ability than by their prior educational opportunities and access to digital infrastructure (Henderson et al., 2017). As a result, treating technological skill gaps as individual deficits rather than structural inequities risks reinforcing existing disadvantages.

Technological confidence also plays an important role in academic success. Technology anxiety can impair performance on digital assessments by increasing cognitive load, even among learners with adequate technical skills (Chua et al., 1999; Compeau & Higgins, 1995). Conversely, instructional approaches that reduce anxiety while building digital competence can improve engagement and academic performance. Culturally responsive pedagogy may be particularly effective in supporting these transitions by explicitly scaffolding learners' technological development rather than assuming digital fluency as a prerequisite for learning.

# Methodology

## Research Design

This study used a mixed-methods design integrating quantitative measures of learner perceptions, technological capability, and academic performance with qualitative open-ended survey data. A concurrent triangulation approach was used, with both data strands collected during the same period and integrated during interpretation to provide a more comprehensive understanding of learner experiences (Creswell & Clark, 2018). The study was conducted in a naturalistic educational setting without random assignment or a control group, prioritizing contextual understanding and ecological validity over experimental control (Lincoln & Guba, 1985).

## Setting and Participants

The study was conducted at a higher education institution in Denmark that offers undergraduate programs specifically designed for international learners. The courses examined were undergraduate marketing courses delivered in English. Participants were undergraduate learners enrolled across the study period, the majority of whom were nationals of Nepal and Bangladesh. The total number of learners surveyed was approximately 105. For the pre-post technology capability analysis, 92 matched pairs were retained, that is, participants for whom complete data were available at both the pre-intervention and post-intervention measurement points. The academic performance analysis was conducted on the subset of learners who completed the course's assessed component.

## The CARE² Model Instructional Intervention

The intervention examined the implementation of the CARE² model in an undergraduate marketing course. The framework organizes culturally responsive instruction around four dimensions: Culture and Collaboration, Agency and Agility, Relevance and Repetition, and Engagement and Evaluation (Fujii & Perez, 2025). Implementation included culturally relevant content, scaffolded digital and assessment tasks, timely formative feedback, and collaborative learning activities. The blended course used Moodle for course materials, LearningZone for assessment submission, and Microsoft PowerPoint for presentations. Supporting learners' use of these technologies was an integral component of the CARE² model approach, reflecting its focus on the intersection of cultural responsiveness and digital equity.

## Data Sources and Measures

Data were collected from four sources: the CARE² model perception survey, pre- and post-intervention technology capability assessments, a knowledge-based course assessment, and final examination performance.

- **CARE² Model Perception Survey.** Learner perceptions were measured using a nine-item, five-point Likert survey (1 = Strongly Disagree to 5 = Strongly Agree) assessing the model's core dimensions: cultural relevance, instructor support, meaningful content, and timely feedback. Descriptive statistics (means and standard deviations) were calculated for each item, the overall scale, and internal consistency was evaluated using Cronbach's alpha.
- **Technology Capability Self-Assessment (Pre/Post).** Technological capability was measured using a 12-item, five-point self-assessment administered before and after the

intervention. Items assessed confidence using course technologies, including PowerPoint, Moodle, LearningZone, online research tools, and ChatGPT. Composite technology scores were calculated by averaging item responses at each time point. Participants also identified their primary technology-related challenges and provided open-ended reflections on their learning experiences.

- **Knowledge-Based Assessment.** Academic performance was measured using scores on Assessment 2 (A2), a course-embedded assessment graded as a percentage with a passing score of 50%. Because A2 was part of the regular course rather than a researcher-developed instrument, it provides an authentic measure of academic achievement.

### Analytical Approach

Quantitative data were analyzed using descriptive and inferential statistics. Pre- and post-intervention technology capability scores were compared using Wilcoxon signed-rank tests, with matched-pairs rank-biserial correlation (r) reported as the effect size. Spearman's rank correlation (ρ) was used to examine the relationship between pre-intervention technology capability and academic performance. Statistical significance was set at $p < .05$. In addition, qualitative responses were analyzed thematically using Braun and Clarke's (2006) six-phase approach. Analysis focused on learners' descriptions of technology-related challenges, confidence, and engagement before and after the intervention. Qualitative findings were integrated with the quantitative results to provide contextual insight into observed statistical patterns.

## Results

The results are presented across four interrelated strands, corresponding to the study's data sources and research questions: learner perceptions of the CARE² model instructional model, pre- and post-intervention technology capability, laptop challenge profiles, and academic performance outcomes including subgroup analyses by country of origin. Qualitative data from open-ended survey responses are integrated throughout to contextualize and deepen the quantitative findings.

### Learner Perceptions of the CARE² Model

Survey responses from 96 learners indicated consistently positive perceptions of the CARE² model instructional approach. The overall mean was 4.07 out of 5.00 (SD = 0.60), with at least 70% of respondents selecting *Agree* or *Strongly Agree* for every survey item. Collaboration (Instructor) (M = 4.24) and Relevance (M = 4.22) received the highest ratings, while Evaluation (M = 3.97) and Culture (M = 3.89) were rated slightly lower but remained favorable. The comparatively lower rating for Culture, a core dimension of the CARE² model framework, is examined further in the discussion. In addition, qualitative responses collected post-intervention reinforced the quantitative perception data, with learners frequently commenting on the quality of instruction and their growing sense of engagement with course material. Representative responses included reflections such as “I love the way you teach us any topic” and “I find it interesting in your class,” suggesting that positive perceptions extend beyond task completion to encompass affective engagement with the learning environment.

## Pre-Post Intervention Change

Matched pre- and post-intervention technology capability data were available for 92 learners. As shown in Table 1, statistically significant improvements were observed in 8 of the 12 technology skills assessed. The largest gains were in LearningZone assessment submission, inserting images into PowerPoint, and creating PowerPoint presentations, all with large effect sizes. Significant improvements were also found for library database use, overall digital skills, Moodle navigation, saving PowerPoint files, familiarity with emerging technologies such as ChatGPT, and internet searching. No significant changes were observed for narrating PowerPoint presentations, private email use, or comfort with learning new technology. See Table 1.

**Table 1**
*Pre/Post Technology Self-Assessment Results*

| Items | Pre Mean | Post Mean | Delta Change | Effect (r) | Significance (p) |
|---|---|---|---|---|---|
| PPT: Make presentation | 3.16 | 4.12 | +0.96 | 0.64 | $p < .05$ |
| PPT: Insert images | 3.55 | 4.73 | +1.18 | 0.77 | $p < .05$ |
| PPT: Narrate | 3.25 | 3.50 | +0.25 | 0.00 | Not significant |
| PPT: Save as .pptx | 3.63 | 4.47 | +0.84 | 0.53 | $p < .05$ |
| Private email use | 3.24 | 3.33 | +0.09 | 0.08 | Not significant |
| Digital skills self-rating | 3.10 | 3.55 | +0.46 | 0.46 | $p < .05$ |
| Comfort learning new tech | 3.46 | 3.60 | +0.14 | 0.12 | Not significant |
| Navigate Moodle | 3.71 | 4.21 | +0.50 | 0.48 | $p < .05$ |
| Submit online (LearningZone) | 2.55 | 3.96 | +1.40 | 0.76 | $p < .05$ |
| Follow new tech (ChatGPT) | 3.33 | 3.57 | +0.24 | 0.23 | $p < .05$ |
| Internet search confidence | 3.39 | 3.61 | +0.22 | 0.21 | $p < .05$ |
| Library database confidence | 2.65 | 3.40 | +0.75 | 0.51 | $p < .05$ |

Notably, no significant improvement was observed for presentation narration despite gains in other presentation skills, nor did comfort with learning new technology change substantially.

These findings suggest the intervention was more effective at developing task-specific digital competencies than at building broader technology confidence. Overall, the CARE² model-informed approach was associated with significant gains across most digital skill areas, particularly those directly related to course assessments, including online submission, presentation development, and digital resource use.

**Laptop Challenge: Shifts in Self-Reported Difficulties**

Complementing the item-level findings, the laptop challenge data showed a marked decline in technology-related anxiety. Learners identifying nervousness or inexperience as their primary challenge decreased from 52 to 24 (54%), while those reporting no challenges increased from 28 to 37. These results align with the technology capability findings, suggesting the intervention was associated with both improved digital skills and reduced technology-related anxiety.

The qualitative data lend texture to this pattern. Pre-intervention responses were characterized by candid expressions of apprehension and unfamiliarity: learners described having "never been using a laptop in class" and feeling "pretty nervous about assessments," with several noting that digital tools such as PowerPoint were ones they were "not able to use perfectly." One learner situated this difficulty explicitly within a cross-national educational context: "As in Nepal we use a different system… not much used to computers." These responses reflect not simply a lack of technical skill but a deeper experience of cultural and pedagogical displacement that the CARE² model framework was designed to address. Post-intervention qualitative responses indicated a clear shift in register, from anxiety and unfamiliarity toward confidence, curiosity, and positive affect. Learners described having "learned to access so many things on the laptop than before" and reported "gaining a lot of knowledge about the laptop and assessment." The provisional quality of some responses, "going pretty smoothly but it will take a bit more to get a good grip," reflects an honest and grounded sense of growth that falls short of mastery, which is arguably more credible and more informative than uniformly enthusiastic responses would be.

Assessment upload anxiety remained a persistent challenge. Although LearningZone submission skills improved significantly, 34 learners reported upload-related difficulties at both measurement points. This difference may reflect the gap between perceived competence in practice settings and anxiety during high-stakes assessment submission. At baseline, 30% of learners preferred paper-based submission and 42% were uncertain, highlighting the significant digital transition faced by the cohort and providing context for both the observed gains and remaining challenges.

**Overall Performance and Technology Capability**

Assessment 2 (A2) grades averaged 55.7% (SD = 12.0), ranging from 31% to 78%. The overall pass rate was 67.6%, with 16.2% of learners achieving A-level grades (≥70%), indicating moderate achievement with substantial individual variation. Spearman's rank correlation showed no significant relationship between pre-intervention technology capability and academic performance ($\rho = 0.02$, ns). This suggests that initial differences in digital skills did not predict academic outcomes, indicating that the CARE² model-informed scaffolding may have helped reduce the impact of prior technological preparedness on learner success.

## Discussion

This study examined the implementation of the CARE² model within an international higher education context characterized by substantial cultural, linguistic, and technological diversity. Across quantitative and qualitative findings, the results suggest that the CARE² model-informed instructional approach was associated with positive learner perceptions, meaningful growth in several forms of technological capability, and the mitigation of inequities that might otherwise have emerged from differential levels of prior digital preparedness. Taken together, the findings contribute to growing evidence that culturally responsive pedagogy in higher education can support not only affective and relational dimensions of learning but also measurable forms of academic and technological adaptation in internationally diverse classrooms.

### Learner Perceptions and the Relational Foundations of CARE² Model

Positive learner perceptions across all CARE² model dimensions suggest that learners experienced the instructional environment as supportive, relevant, and engaging. The strongest ratings for instructor collaboration and content relevance align with research on culturally responsive pedagogy that emphasizes relational trust, instructor accessibility, and meaningful learning connections (Gay, 2000; Ladson-Billings, 1995). In international contexts, these factors may be especially important as learners navigate unfamiliar linguistic, technological, and academic systems. The high relevance rating further suggests that learners perceived course activities as connected to their experiences and goals, supporting engagement despite initial technological challenges (Hammond, 2015).

The Culture dimension received the lowest rating, although it was still evaluated positively. This finding may suggest that while learners experienced strong support and relevance, deeper integration of cultural identities and perspectives was less visible. Culturally responsive pedagogy requires more than accommodation; it involves intentional attention to cultural assumptions within curriculum, communication, participation, and assessment practices (Gay, 2018; Hockings, 2010). In transnational settings, learners may distinguish between feeling supported within an existing system and seeing their cultural knowledge meaningfully represented within that system. This distinction highlights the need for continued refinement of culturally responsive frameworks in international higher education.

### Technological Capability Growth and Digital Equity

A key finding was the substantial improvement in technological capability, with significant gains in eight of twelve assessed skill areas. The largest improvements occurred in competencies directly related to course participation and assessment, including online submission, PowerPoint development, image insertion, and learning management system navigation. These findings suggest that digital preparedness should be viewed as a developmental capacity that can be supported through responsive instructional design rather than a fixed prerequisite for academic success.

This finding contributes to digital equity research by challenging deficit-based interpretations of technological gaps among international learners (Bond et al., 2021; Czerniewicz et al., 2020). Although many learners entered the course with limited experience in digitally mediated learning environments, structured support enabled meaningful growth, reinforcing the view

that technology disparities are often shaped by educational context and access rather than individual ability (Henderson et al., 2017).

Qualitative responses further indicated that learners' challenges extended beyond technical skills to include uncertainty and anxiety associated with unfamiliar academic technologies and practices. This suggests that digital adaptation in international higher education involves navigating new technological, cultural, and institutional expectations simultaneously. The reduction in reported nervousness and inexperience also indicates that the intervention may have reduced technology-related anxiety, which can independently affect performance by increasing cognitive load (Compeau & Higgins, 1995). Through scaffolding, repeated exposure, and normalization of uncertainty, culturally responsive instruction may support both digital competence and learner confidence.

### Persistence of Assessment Submission Anxiety

Despite overall gains, persistent anxiety surrounding formal assessment uploads remains an important finding. Although learners reported improved competence with online submission systems, many continued to experience upload-related concerns, highlighting a distinction between technical ability and confidence under high-stakes conditions. For international learners navigating unfamiliar institutional systems, academic expectations, and digital assessment practices, confidence is shaped not only by procedural skill but also by trust, predictability, and perceived risk. These findings suggest that digital submission should be treated as a learning process rather than an administrative task. Institutions may benefit from repeated, embedded practice opportunities that simulate authentic assessment conditions. The CARE² model principles of Repetition and Evaluation are particularly relevant, as ongoing guided exposure may reduce anxiety associated with high-stakes digital tasks.

### Technology Preparedness and Academic Equity

Perhaps the most theoretically significant finding was the absence of a meaningful relationship between pre-intervention technology capability and academic performance. Learners with lower initial digital confidence did not perform worse than peers with stronger technological backgrounds, suggesting that prior digital preparedness did not determine academic outcomes when appropriate scaffolding was provided. These finding challenges deficit-based views of digital readiness and highlights the equity potential of culturally responsive instructional design. Rather than requiring learners to arrive with institutionally valued technological skills, the CARE² model-informed approach supported learners in developing these competencies during the course, reducing the extent to which prior access and experience shaped academic success. More broadly, the findings suggest that equitable internationalization requires not only diverse enrollment but also instructional structures that enable meaningful participation.

### Implications for International Higher Education

Several implications were identified from these results for institutions serving internationally diverse learners. First, culturally responsive pedagogy may extend beyond its traditional K-12 applications by providing a framework for addressing the cultural and technological transitions experienced by international university learners. The findings suggest that such approaches can support both learner engagement and the development of digital competencies required for academic participation. Second, digital equity should be embedded within curriculum design rather than treated as a separate support concern. Equitable instruction requires intentional

scaffolding of digital practices, expectations, and confidence, not simply access to technology. Finally, internationalization should be viewed not only as demographic diversification but as pedagogical transformation requiring instructional systems that respond to differences in learners' educational backgrounds, technological experiences, and cultural approaches to learning.

## Limitations and Future Research

This study has several limitations. The non-experimental design and the absence of a control group limit causal claims, as observed improvements may reflect maturation, increased exposure to technology, or other contextual factors. Self-reported technology measures may also be influenced by subjective interpretation, suggesting future research should incorporate performance-based assessments of digital competence. Additionally, the single institutional and disciplinary context limits generalizability. Future research should examine the CARE² model framework across disciplines, institutions, and cultural contexts using longitudinal and comparative designs. Further investigation of the Culture dimension is particularly warranted to understand how learners experience cultural responsiveness and how cultural integration, representation, communication, and assessment practices influence learners' sense of belonging and achievement.

## Conclusion

The findings of this study suggest that the CARE² model is a promising framework for supporting international learners as they navigate culturally and technologically unfamiliar higher education environments. Learners reported positive perceptions of the instructional environment, demonstrated meaningful growth across multiple domains of technological capability, and achieved academic outcomes that were not systematically constrained by their initial levels of digital preparedness. Together, these findings support the argument that culturally responsive pedagogy in higher education can serve not only as an inclusive teaching philosophy but also as a mechanism for promoting digital equity and reducing structural barriers to participation. In increasingly internationalized universities, the challenge is no longer simply whether institutions can recruit globally diverse learners, but whether they can design learning environments in which those learners can participate fully and equitably. The present findings suggest that culturally responsive and technologically scaffolded instructional design may be an important part of that answer.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors used Grammarly for proofreading and language refinement, and OpenAI to simplify research questions for presentation consistency. Qualitative analyses involved inter-coder review, with themes finalized through coder agreement and supplemented by automated validation using OpenAI 5o. Quantitative analyses included assessment data and post-unit surveys measuring learning outcomes and learner perceptions. No AI or AI-assisted tools were used to generate manuscript content; all research design, procedures, findings, analyses, and interpretations are original contributions of the authors.

## References

Bond, M., Bedenlier, S., Marín, V. I., & Händel, M. (2021). Emergency remote teaching in higher education: Mapping the first global online semester. *International Journal of Educational Technology in Higher Education, 18*(1), 50. https://doi.org/10.1186/s41239-021-00282-x

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education, 31*(2), 219–233. https://doi.org/10.1080/03075070600572132

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Chua, S. L., Chen, D. T., & Wong, A. F. L. (1999). Computer anxiety and its correlates: A meta-analysis. *Computers in Human Behavior, 15*(5), 609–623. https://doi.org/10.1016/S0747-5632(99)00039-4

Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly, 19*(2), 189–211. https://doi.org/10.2307/249688

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Czerniewicz, L., Agherdien, N., Badenhorst, J., Belluigi, D., Chambers, T., Chili, M., de Villiers, M., Felix, A., Gachago, D., Gokhale, C., Ivala, E., Kramm, A., Madiba, M., Mgqwashu, E., Pallitt, N., Prinsloo, P., Solomon, K., van der Merwe, T., & Wissing, G. (2020). A wake-up call: Equity, inequality and COVID-19 emergency remote teaching and learning. *Postdigital Science and Education, 2*(3), 946–967. https://doi.org/10.1007/s42438-020-00187-4

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109.

Fujii, K. K., & Perez, N. (2025). The CARE$^2$ model: an action research study exploring multicultural experiences in an international higher education classroom. *International Journal for Educational Media and Technology, 19*(2). https://www.ijemt.org/index.php/journal/article/view/382

Fujii, K. K. & Perez, N. (2026). The CARE$^2$ model: a practitioner-oriented framework to support instructional designers build culturally responsive learning experiences. *The Journal of Applied Instructional Design, 15*(3). https://doi.org/10.59668/2580.24743

Gay, G. (2018). *Culturally responsive teaching: Theory, research, and practice* (3rd ed.). Teachers College Press.

Gay, G. (2000). *Culturally responsive teaching: Theory, research, and practice*. Teachers College Press.

Hammond, Z. (2015). *Culturally responsive teaching and the brain: Promoting authentic engagement and rigor among culturally and linguistically diverse students*. Corwin.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Henderson, M., Selwyn, N., Finger, G., & Aston, R. (2017). Students' everyday engagement with digital technology in university: Exploring patterns of use and 'usefulness'. *Journal of Higher Education Policy and Management, 39*(3), 308–319.

Hockings, C. (2010). *Inclusive learning and teaching in higher education: A synthesis of research*. Higher Education Academy.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38*(5), 365–379. https://doi.org/10.3102/0013189X09339057

Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of Instructional Development, 10*(3), 2–10. https://doi.org/10.1007/BF02905780

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465–491. https://doi.org/10.3102/00028312032003465

Leask, B. (2015). *Internationalizing the curriculum*. Routledge.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.

Montgomery, C. (2010). *Understanding the international student experience*. Palgrave Macmillan.

Mthethwa-Sommers, S. (2014). *Narratives of social justice educators: Standing firm*. Springer.

Reeve, J. (2012). A self-determination theory perspective on student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 149–172). Springer. https://doi.org/10.1007/978-1-4614-2018-7_7

Ryan, J. (2011). *Teaching and learning for international students: Towards a transcultural approach*. Routledge.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Sleeter, C. E. (2012). Confronting the marginalization of culturally responsive pedagogy. *Urban Education, 47*(3), 562–584. https://doi.org/10.1177/0042085911431472

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

# Process-Oriented Collaborative Essays: Empowering Undergraduate Students Through Assessment Redesign in the Age of Generative AI

Fahima Alam, Queen Mary University of London, United Kingdom
Jonathan Jackson, Queen Mary University of London, United Kingdom

The European Conference on Education 2026
Official Conference Proceedings

## Abstract

The growing use of Generative Artificial Intelligence (GenAI) tools such as ChatGPT is reshaping how students approach learning and assessment in Higher Education (HE) as well as the viability of traditional assessment methods such as essays. While a significant amount of research has focused on student acceptance of GenAI through models such as the Technology Acceptance Model (TAM) and Unified Theory of Acceptance and Use of Technology (UTAUT), there is a pressing need for practical examples of how assessments in HE can remain valid and secure despite the prevalence of GenAI. This study explores the benefits and limitations of a collaborative, process-oriented assessment implemented with 23 first-year degree apprenticeship students at Queen Mary University of London. Student teams were required to engage with an iterative essay-writing task, producing multiple drafts and reflecting on how their work developed through collaboration with each other and optional use of GenAI tools to support their writing process. Drawing on Task-Technology Fit (TTF) and Institutional Theory, the study examines how GenAI aligns with academic writing practices and institutional expectations within an undergraduate assessment context. The findings suggest that GenAI can support the learning process through exploration of ideas and refinement of arguments over time, if appropriate assessment design scaffolding exists. The iterative and collaborative elements of the assessment appeared to support the development of students' academic voice, identity, and autonomy. The study highlights the value of designing assessments that prioritise process, reflection, and collaboration, enabling more meaningful and engaging learning experiences while strengthening academic integrity.

*Keywords*: GenAI, process-oriented assessment, assessment design, higher education

iafor
The International Academic Forum
www.iafor.org

## Introduction

In recent years, the emergence of Generative Artificial Intelligence (GenAI) tools has disrupted the higher education (HE) landscape, with the conversation shifting from experimentation towards wider discussion on how these tools influence learning, teaching, and assessment (Cotton et al., 2024). Many arguments have centred on academic integrity, specifically the possibility of students submitting work generated in part or entirely by GenAI and the difficulties associated with detecting these practices (Kofinas et al., 2025; Perkins, 2023). Specific Large Language Model (LLM) chatbots such as ChatGPT have gained significant attention as they are able to generate ideas, propose structures, summarise information, refine arguments and draft text (Kasneci et al., 2023). Student use is also widespread with a 2026 study conducted by the Higher Education Policy Institute (HEPI) revealing that 94% of the students surveyed had used GenAI in some way for assessed work (Stephenson & Armstrong, 2026). Therefore, the issue is not about whether or how GenAI might be accepted. Rather, a more complex question must be asked: how should assessments be designed when emerging digital technologies can perform some of the required tasks that students have historically been expected to complete?

Bouaziz and Eddouada (2026) state that essay-based assessments are widely used as a method of evaluating the subject knowledge, critical thinking, and the analytical writing capabilities of the students. Nonetheless, the final essay only offers a partial representation or proxy of the intellectual processes that should be involved in essay writing, providing limited understanding on how students have built ideas, engaged with feedback, and revised their arguments over time (Lee, 2006; Mueller-Lyaskovets & Horner, 2021). As GenAI has become almost ubiquitous and capable of supporting and mimicking various aspects of academic writing, it is challenging to ascertain the extent to which a final written artefact reflects the knowledge, understanding, and intellectual development of the student (Ali & Maroulis, 2026). But this does not diminish the educational value of essay writing; instead, it raises the question of whether the final written essay, when considered in isolation, is able to offer sufficient evidence of the learning that an essay assessment is intended to measure. Cheng et al. (2024) assert that it is important to be attentive towards assessment approaches that make the reasoning, decision-making, and the development of the student more visible throughout the learning journey. It is with this principle in mind, that the work outlined in this paper was undertaken.

## Significance of the Study

Many early studies around GenAI and assessment have emphasised detection, academic integrity, and the likelihood of students submitting work generated by GenAI as their own (Cotton et al., 2024; Perkins, 2023). As an emerging challenge and a wider pedagogical question, educators must ask what assessment approaches are most supportive of evidencing student learning, critical thinking and intellectual development (Ardito, 2025). Furze et al. (2024) argued that blocking or banning GenAI is unlikely to offer a sustainable response but instead illustrate how assessment might be redesigned across various levels of appropriate AI engagement. In a similar vein, Bannister et al. (2025) assert that it is important to assess the validity and susceptibility of current assessments within an environment where GenAI tools are widely available. Furthermore, systematic reviews of assessment within AI-infused learning environments have discovered the requirement to reconsider traditional assessment strategies to preserve meaningful assurance of learning (Al-Abdullatif & Alsubaie, 2024).

This paper presents one practical response to this challenge: a collaborative process-oriented essay assessment designed for first year undergraduate students. This paper also draws on two complementary theoretical perspectives: Task–Technology Fit (TTF) and Institutional Theory. TTF is used to examine the extent to which the capabilities of GenAI align with the requirements of specific academic writing tasks, while Institutional Theory provides a lens for considering whether such uses are consistent with institutional expectations surrounding academic integrity, authorship and legitimate evidence of student learning. Bringing these perspectives together facilitates the consideration of not only whether GenAI can effectively support an academic task, but also whether its use remains educationally and institutionally appropriate within the assessment context.

The core argument of this paper is that the value of process-oriented assessment does not lie in attempting to make assessment resistant to GenAI. Rather, it lies in making students' intellectual contribution more visible by providing evidence of how they develop ideas, exercise judgement, respond to feedback, collaborate with others and refine their work over time.

## Literature Review

### Impact of GenAI on Assessment in HE

Across the HE landscape, GenAI has become a prominent topic of discussion, arguably creating new opportunities for learning but at the same time, posing challenges for many extant approaches to assessment (Cotton et al., 2024). Kasneci et al. (2023) state LLMs can support many worthwhile activities such as personalised learning, idea generation, providing constructive feedback and other aspects of academic writing while acknowledging concerns around bias, misinformation, overdependence, and the extent to which students remain intellectually accountable for the work they generate. Hence, GenAI should not be comprehended as either inherently effective or detrimental to learning, but in the context of education it is important to understand how it could be employed in academic tasks and how students might continue to use it while maintaining critical judgement over generated outputs.

While there are legitimate reasons for some entirely AI-free learning environments, this has become an increasingly difficult proposition for HE to sustain. As noted in the introduction of this paper, as of 2026, 94% of students surveyed as part of HEPI research have used GenAI in some way for assessed work (Stephenson & Armstrong, 2026) so this cannot be treated lightly. As students are increasingly likely to encounter AI-supported practices within future professional environments, completely excluding GenAI from learning and assessment may restrict opportunities to develop the critical and responsible AI literacy required beyond higher education (Upsher et al., 2024; Su et al., 2026). Similarly, it is the responsibility of universities to ensure that assessment outcomes provide credible evidence of what students know and what they can do. This combination of competing demands makes GenAI a "wicked problem" for assessment (Corbin et al., 2025) in that there is no single correct solution to the problem. In some contexts, the prohibition of GenAI might protect certain conceptions of independent authorship but may minimise authenticity. On the other hand, unrestricted GenAI use might undermine individual students' ability to learn and be able to establish themselves in contemporary and future workplaces.

Thus, the challenge for HE extends beyond deciding whether GenAI should be prohibited or permitted. Fundamentally, a re-consideration is needed of how assessment can accommodate

increasingly AI-mediated learning while continuing to offer meaningful and valid evidence of knowledge, judgement, and intellectual development of the student.

## Assessment Validity & Academic Integrity in the Age of GenAI

Within HE, GenAI has sparked two important concerns: assessment validity and academic integrity. From the viewpoint of academic integrity, attention has been concentrated on whether students are the main authors of an essay, for example, and whether the usage of GenAI has been consistent with the expectations of the institution in the context of appropriate academic practice (Cotton et al., 2024; Kofinas et al., 2025). Essentially, has the student remained intellectually accountable for their work or have they offloaded it to an extent which is deemed academic misconduct?

On the other hand, assessment validity raises different concerns such as whether the assessment facilitates the production of suitable evidence by students that they have the knowledge and skills that the assessment was anticipated to measure. Of course, academic integrity and assessment validity can overlap but they should not be conflated. For instance, a student can make use of GenAI within the institutional boundaries and therefore satisfy the academic integrity requirements, but the resulting assessment might offer limited suitable evidence of a particular capability or competency.

Importantly, the use of GenAI does not negate the assessment validity where students are still accountable for assessing the output, making decisions, and building their own argument. This is validated by two studies which emphasise the significance of considering the relationship between GenAI usage and what the assessment is intended to evidence (Bannister et al., 2025; Sharma, 2026).

In fact, Sharma (2026) suggests that the student's judgement can be more visible when the assessment is supported by GenAI. The quality of the final submission is significant, but this can be reinforced by supplementary evidence that shows how the student achieved the outcome. Essentially, process-oriented assessment can help in understanding the reasoning, decisions, and development of the students' knowledge and at the same time evaluate the final written work.

## Process-Oriented Assessments

Process-oriented assessment offers a possible response to the various challenges surrounding opaque written assessments in the age of GenAI. Conventionally, an essay assessment is focused on the quality of the final submission where the final written essay acts as the sole piece of evidence for evaluating the performance of the student (Lee, 2006). Process-based writing literature has argued that writing should be comprehended as an activity of development that involves drafting, planning, feedback, revision, and reflection instead of solely focusing on the finished product (Lee, 2006; Mueller-Lyaskovets & Horner, 2021). Thus, a process-oriented assessment helps in broadening the evidence that is available by making the chosen stages of a student's development more explicit.

There is an important distinction between accepting a suggestion generated by GenAI without questioning it and evaluating the suggestion against academic evidence, discussing it with others, and then deciding whether to reject or accept it. Cheng et al., (2024) conducted research examining human-AI collaborative writing and they similarly recommend that process

evidence can offer information about how students are interacting with GenAI that might not be visible from the completed text alone.

Nonetheless, simply asking students to provide several drafts of their work does not necessarily offer strong evidence of learning. It is important for the process-oriented assessments to emphasise the purpose and quality of development instead of the number of documented stages (Lee, 2006). All too easily, an assessment designed with good intentions could lead to students making relatively trivial changes between drafts simply to satisfy the requirement of showing multiple revisions, or they could concentrate on performatively creating process-oriented artefacts rather than engaging in the process in a way that develops their thinking. The value of process evidence therefore lies less in showing that revisions took place (which risks a performative response) and more in revealing the judgement behind those revisions. For example, students might explain why an argument was reconsidered, how evidence changed their position, how peer feedback influenced a decision, or why a GenAI-generated suggestion was accepted, modified or rejected.

This approach also connects with emerging discussions of authentic assessment in relation to the pervasive presence of GenAI. Su et al. (2026) argue that assessment needs to recognise the evolving role of GenAI while continuing to create opportunities for students to demonstrate meaningful learning. Similarly, frameworks for responsible GenAI use emphasise that different levels of AI involvement should be considered in relation to the purpose and intended learning outcomes of an assessment (Ucan & Demirel Ucan, 2026). The transition away from product-oriented assessment towards process-oriented assessment should therefore not be understood as abandoning the final essay. Instead, the final artefact can remain an important indicator of achievement while being complemented by evidence of how students developed it (Midford, 2025). In this way, iteration, collaboration and reflection can provide additional insight into students' reasoning, decision-making and intellectual contribution (Cheng et al., 2024).

## Collaboration, Reflection and Academic Voice

Collaboration is seen to be pedagogically important as it facilitates important human-to-human interaction with the intention of mitigating risk of student isolation. Also, collaborative writing requires students to respond to different interpretations, negotiate meaning, justify choices and coordinate the development of a shared text. Scaffolding these activities in class provides visible social learning and can offer a supplementary source of responsibility as decisions are exposed to peers instead of working privately (Gündüz, 2023; Khuder & Negretti, 2026).

Additionally, reflection can provide another layer to the assessment. The concept of reflection-on-action (Schon, 1983) is useful for helping students understand how they might retrospectively assess decisions and experiences. Within an assessment which utilises GenAI, reflective documentation can help in revealing not only what students have changed but why they have changed them. For example, providing an explanation for why an AI generated suggestion was rejected offers evidence of evaluative judgement that would be absent from the final artefact. Therefore, reflection can operate simultaneously as a learning activity as well as evidence for process-oriented assessment.

This paper proposes that the combination of collaboration and reflection together are particularly valuable within the scope of process-oriented assessment. As mentioned previously, collaboration can create opportunities for ideas to be negotiated, questioned, and refined through interaction with others whereas reflection can offer students to explain the

reasoning behind the decisions. These approaches transition responsibility away from relatively ineffective declarations of GenAI use towards demonstrating how students exercise judgement over its use. This approach seems to be consistent with emerging assessment frameworks encompassing GenAI that suggest aligning the permitted usage of AI with intended learning outcomes and at the same time retaining meaningful opportunities for critical thinking and human input (Furze et al., 2024; Perkins et al., 2024).

## Theoretical Frameworks

### Technology-Task Fit (TTF)

TTF considers whether the capabilities of a technology align with the requirements of a particular task. In the context of this study, it is helpful to examine whether GenAI genuinely supports different aspects of academic writing or whether the fit is more limited. In their review of TTF literature, Spies et al. (2020) showed that the TTF framework has been applied across a wide range of technology-use contexts. In relation to GenAI, TTF is useful because it moves the discussion away from simply asking whether AI is beneficial or harmful to academic writing. Instead, it allows us to ask a more specific question: how well do the capabilities of GenAI fit different academic writing tasks?

More recent research further supports the relevance of TTF to GenAI. Wang et al. (2026) examined the relationship between TTF, perceived usefulness and responsible GenAI use among 280 university students. They found that a stronger fit between academic tasks and GenAI capabilities was associated with greater perceived usefulness and more responsible use. The study also extends established technology-adoption approaches, such as the Technology Acceptance Model (TAM), by considering not only whether students find a technology useful, but also how responsibly it is used. This is particularly relevant to assessment, where the concern is not simply whether students use GenAI, but how its use relates to the intended learning outcomes.

In academic writing, for example, GenAI may have a strong task fit for generating possible counterarguments, exploring alternative structures, brainstorming ideas or identifying language issues. However, the fit becomes more complex when students are expected to demonstrate their own judgement, evaluate the quality of evidence, apply professional experience or negotiate different viewpoints with peers. TTF therefore allows GenAI to be considered in relation to specific academic writing activities, rather than treating the essay as one single task.

### Institutional Theory

Institutional Theory helps in the understanding of expectations and contextual norms within which universities operate. Assessment does not take place in isolation; it is influenced by institutional policies, regulatory requirements and established expectations about what counts as evidence of student learning (Cai & Mehari, 2015). This means that assessment design is not only shaped by what works well for teaching and learning. Universities must also consider issues such as authorship, fairness, academic integrity, quality assurance and the credibility of the qualifications they award.

Using Institutional Theory alongside TTF facilitates a consideration of both task fit and institutional fit. For example, GenAI may be very effective at generating or improving written text, meaning that the technology can fit certain academic writing tasks well. However, this

does not necessarily mean that its use is appropriate for every assessment. If learning outcomes require students to develop, evaluate and communicate their own argument, extensive reliance on GenAI may make it more difficult to interpret the student's work without any evidence of their process. At the same time, allowing appropriate use of GenAI may be valuable where developing AI literacy is relevant to the learning outcomes and students' future professional practice. Therefore, the question is not simply whether GenAI can support a particular task, but whether the way it is used also fits the purpose of the assessment and the expectations of the institution.

## Methodology

Rather than a controlled research intervention, the work outlined in this paper was undertaken as part of programme-wide pedagogical redesign. This study therefore adopts a practice based exploratory perspective and focusses on the resulting observations as emerging insights instead of generalisable empirical findings.

The collaborative process-oriented essay assessment was designed for first year Business Analyst degree apprenticeship students at Queen Mary University of London. The cohort size was 23 students, made up of 11 self-selecting teams, one of which was a team of 3.

Students were required to develop a 1500-word essay in their teams and document at least two stages of iterative development. Each team was required to choose from three pre-determined essay titles, which related to the main topic of the module. The assessment criteria made clear how evidence of iteration and collaboration would contribute to the final grade, as demonstrated by Table 1. While by no means perfect, these assessment criteria ensured that the assessment was structurally designed to require evidence of these activities (Corbin, Dawson, et al., 2025).

Students were permitted to GenAI within clear boundaries, but use of GenAI was not mandatory. The following Generative AI guidance was included in the assessment brief for clarity and to encourage good academic practice and avoid common pitfalls:

> Use of Generative AI (e.g., MS Copilot or ChatGPT) is permitted for this assessment but must not be used to generate text which you then submit as your own as this is academic misconduct. If using spelling or grammar assistive tools, avoid using any Generative AI functionality which rewrites non-trivial portions of your work. Avoid relying on Generative AI to format your references as these tools can often produce references which do not exist which will significantly undermine the credibility of your work and may be considered academic misconduct.

While this statement can be viewed as discursive rather than structural, it was still deemed necessary for the purposes of making the expectations of the assessment clear and to provide reasons for how inappropriate use of GenAI can lead to poor outcomes.

**Table 1**
*Sample Assessment Criteria Related to Collaboration and Iteration*

| Criteria | 0-29 (F) | 30-39 (E) | 40-49 (D) | 50-59 (C) | 60-69 (B) | 70-79 (A) | 80-100 (A+) |
|---|---|---|---|---|---|---|---|
| **Collaborative Contribution & Knowledge Co-construction (20%)** | No meaningful evidence of collaboration. | Minimal collaboration; weak evidence of joint working. | Limited evidence of collaboration; engagement appears inconsistent. | Basic collaboration evident but contributions may be uneven. | Effective collaboration with some evidence of shared contribution. | Very effective collaboration with clear evidence of peer engagement and joint responsibility. | Exceptional evidence of balanced collaboration, shared decision-making, and collective knowledge construction throughout the task. |
| **Iterative Development & Evidence of Revision (20%)** | No evidence of iterative development. | Drafts submitted but show negligible revision or learning progression. | Minimal evidence of iteration; little meaningful development between drafts. | Some iteration evident; changes are mostly surface level with minimal explanation. | Evidence of at least 2 iterations with noticeable improvement, though justifications may lack depth. | Strong evidence of at least 2 meaningful iterations with clear improvement and rationale for revisions. | In addition to A-grade requirement: Evidence of insightful reflections on the collaborative process and shared learning. |

## Discussion

This section outlines some of the co-authors' insights and reflections after administering the process-oriented collaborative essay assessment for the first time in early 2026.

### GenAI as a Source for Interrogation Instead of Offloading

One noteworthy feature that was evident in some of the submissions was that students who used GenAI queried the suggestions instead of simply accepting them. For example, in some cases the interaction of counterarguments was followed by revisions, active challenging, or rejection. This pattern is vital as it helps in demonstrating that usage of GenAI can indeed incorporate and even stimulate intellectual work instead of replacing it. Therefore, AI can be seen as a dialogic resource that helps stimulate further thought (Miao & Holmes, 2023).

Within the context of TTF, this represents an alignment between task and technology where GenAI is appropriate in rapidly producing alternate positions, however the educational value is based on whether students are being held accountable and sufficiently motivated to assess those positions. In a similar vein, technological capabilities can help in supporting learning or they can undermine it and this is based on the subsequent activities of the student. Therefore, the assessment design should emphasise not only whether GenAI is permitted but it should offer structural requirements on how positive interactions with GenAI tools can be sufficiently evidenced.

### Collaborative Learning as Process Evidence

The 11 student teams displayed various approaches to collaboration during the module as they worked on their assessment. Students shared their drafts, exchanged peer feedback, and revised the work in an iterative manner. Some also leveraged Microsoft Teams video call and chat

discussions that offered additional opportunities for them to informally present and explain the reasoning behind specific decisions. Therefore, collaboration offered a social dimension in the process-oriented assessment. All these aspects are important as collaborative writing is not simply about dividing a task between team members. This is evident in recent research that illustrates effective collaboration can involve negotiation, collective decision-making, shared goal setting, and regulation of the developing text (Khuder & Negretti, 2026).

Khuder and Negretti (2026) discovered that students who developed socially shared regulation discussed issues collectively, negotiated various feedback and made decisions together. Additionally, they further advocate that iterative processes are useful as students can justify decisions and can offer clarification which can help in strengthening the collaborative regulation. Therefore, this offers a significant insight for the current study where students provided an explanation to their peers, responded to disagreements, justified their revisions, and offered explanations that may be less visible in individual writing.

In the current era of GenAI ubiquity, the collaborative dimension is important. Some recent work that has integrated peer feedback and GenAI into academic writing recommends that human peer interaction can help in complementing writing that is supported by AI by offering opportunities for critical engagement (Orvalho & Monteiro, 2025). Thus, collaboration can offer an extra layer of analysis around the development where students are required to explain and negotiate choices instead of making them in isolation.

Even though collaboration can offer opportunities, it should not be treated as singular evidence of equitable learning. Likewise, working in a group does not necessarily mean that meaningful collaboration takes place. According to Khuder and Negretti (2026), collaboration is stronger when students can share their goals, discuss decisions, and take collective accountability for the work. This is significant to the current study where 11 teams demonstrated various approaches to collaboration and revision, some more effective than others. Consequently, the authors note that the quality of collaboration could be improved with more scaffolded collaborative activities in class to help students experience different ways of collaborating so that they can then leverage these approaches more autonomously for their collaborative assessment.

**Limitations of Process-Oriented Documentation**

Some students informally mentioned that the required process-based documentation seemed like a lengthy and time-consuming process. Also, some students were not initially sure how to present a meaningful commentary on their process. These lessons are important to consider as it demonstrates how the purpose behind the process-based evidence needs to be made clear to students and that more examples may be beneficial, particularly for students relatively new to university-level education.

Another limitation of process-oriented assessment is the risk of “empty performativity” where apparent evidence produced by students does not reflect any genuine intellectual development. When students understand that collaboration, revision, and reflection are being assessed, they may lean towards producing evidence of activities that meet the perceived expectations of the marker. In a systematic review by Ross et al., (2024) on reflective assessment, it was discovered the expectations of the assessment can encourage performative instead of genuine reflection. Hence, reflections, notes, and records of collaboration should not automatically be viewed as satisfactory evidence of learning.

## Conclusion

The response to GenAI by HE must move beyond attempts to detect whether students have used GenAI tools as part of an assessment. In this paper, collaborative and process-oriented assessment is shown to afford opportunities to make students' learning visible through peer interaction, iterative drafting, revision and reflection. All these elements contribute to a multi-faceted body of evidence which can demonstrate how students are developing arguments, building judgement during the writing process, and responding to feedback. Nonetheless, process-oriented assessments should not be viewed as a solution suitable for all contexts. The differences in participation across groups and the additional workload (perceived or actual) related to documenting an iterative process are factors which can still undermine the validity of the assessment without careful scaffolding and student support.

The authors of this paper plan to strengthen future assessments by communicating more clearly to students the purposefulness of evidencing their process while minimising less valuable documentation. Other emerging frameworks will also be explored such as the Voice-First Written Assessment (VFWA) approach proposed by Webb-Davies et al. (2026) which may offer a beneficial extension to process-oriented essay assessments. VFWA involves capturing "evidence of [students'] task-related reasoning under observed conditions before generating a final written submission" in an environment where GenAI or other assistive digital technologies are available. The approach looks to establish a clear association between the final written work and the reasoning of the student instead of attempting to eliminate the use of technology in writing, which aligns with the underlying principles of the work outlined in this paper. Future research would be able to examine whether the combination of these approaches leads to improved quality of evidence of student judgement while remaining manageable for students and those assessing the work.

In summary, the work presented in this paper suggests that a combination of scaffolded collaborative work and process-oriented assessment design can lead to conditions where reasoning, judgement, and iterative development can remain more visible and offer valuable evidence upon which evaluation of students' work can be based.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The authors declare that ChatGPT was used in refining some language used in the first draft of the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. Subsequently, a full human review and edit was undertaken prior to submission of the manuscript for consideration in the conference proceedings. The authors further declare that, apart from ChatGPT, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of practice-based research.

## References

Al-Abdullatif, A. M., & Alsubaie, M. A. (2024). ChatGPT in Learning: Assessing Students' Use Intentions through the Lens of Perceived Value and the Influence of AI Literacy. *Behavioral Sciences, 14*(9), 845. https://doi.org/10.3390/bs14090845

Ali, R., & Maroulis, J. (2026). Assessment validity in the age of generative artificial intelligence: A critical review. *Assessment & Evaluation in Higher Education, 0*(0), 1–16. https://doi.org/10.1080/02602938.2026.2683046

Ardito, C. G. (2025). Generative AI detection in higher education assessments. *New Directions for Teaching and Learning, 2025*(182), 11–28. https://doi.org/10.1002/tl.20624

Bannister, P., Santamaría Urbieta, A., & Brufau Alvira, N. (2025). Appraising higher education assessment validity: Development of the PANDORA GenAI Susceptibility Rubric. *Journal of Applied Learning & Teaching, 8*(1). https://doi.org/10.37074/jalt.2025.8.1.20

Bouaziz, K., & Eddouada, S. (2026). AI-Assisted Academic Writing and Publishing in Morocco: Doctoral Students' Ethical Decision-Making Accounts. *Arab World English Journal*, (3), 268–286. https://doi.org/10.24093/awej/A3.17

Cai, Y., & Mehari, Y. (2015). The Use of Institutional Theory in Higher Education Research. In J. Huisman & M. Tight (Eds), *Theory and Method in Higher Education Research* (Vol. 1, pp. 1–25). Emerald Group Publishing Limited. https://doi.org/10.1108/S2056-375220150000001001

Cheng, Y., Lyons, K., Chen, G., Gašević, D., & Swiecki, Z. (2024). Evidence-centered Assessment for Writing with Generative AI. *Proceedings of the 14th Learning Analytics and Knowledge Conference*, 178–188. https://doi.org/10.1145/3636555.3636866

Corbin, T., Bearman, M., Boud, D., & Dawson, P. (2025). The wicked problem of AI and assessment. *Assessment & Evaluation in Higher Education, 51*(4), 736–752. https://doi.org/10.1080/02602938.2025.2553340

Corbin, T., Dawson, P., & Liu, D. (2025). Talk is cheap: Why structural assessment changes are needed for a time of GenAI. *Assessment & Evaluation in Higher Education, 50*(7), 1087–1097. https://doi.org/10.1080/02602938.2025.2503964

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Furze, L., Perkins, M., Roe, J., & MacVaugh, J. (2024). The AI Assessment Scale (AIAS) in action: A pilot implementation of GenAI-supported assessment. *Australasian Journal of Educational Technology*. https://doi.org/10.14742/ajet.9434

Gündüz, Z. E. (2023). Teaching Writing with Wiki-Based Collaborative Writing Tasks in an EFL Context at Higher Education. *Acuity: Journal of English Language Pedagogy, Literature and Culture, 8*(1), 68–86.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Khuder, B., & Negretti, R. (2026). Collaborative writing regulation: A comparative case study of co-regulation and socially shared regulation in higher education. *Higher Education, 91*(2), 677–699. https://doi.org/10.1007/s10734-025-01437-9

Kofinas, A. K., Tsay, C. H., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology, 56*(6), 2522–2549. https://doi.org/10.1111/bjet.13585

Lee, Y.-J. (2006). The process-oriented ESL writing assessment: Promises and challenges. *Journal of Second Language Writing, 15*(4), 307–330. https://doi.org/10.1016/j.jslw.2006.09.003

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Midford, S. (2025). Is the essay history? Rethinking effective assessment in the age of generative AI. History Australia, 22(4), 563–567. https://doi.org/10.1080/14490854.2025.2570457

Mueller-Lyaskovets, T., & Horner, O. (2021). Integrating Formative Assessment with Foreign Language (English) Process Writing Instruction: Lessons from Two College Writing and Reading Classes in Germany. *Journal of Academic Writing, 11*(1), 62–79. https://doi.org/10.18552/joaw.v11i1.499

Orvalho, L., & Monteiro, A. F. (2025). *IA Generativa e Feedback entre Pares na Escrita Académica* [Generative AI and peer feedback in collaborative academic writing]. Saber e Educar, 34(1). https://doi.org/10.25767/SE.V34I1.42564

Perkins, M. (2023). Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice, 20*(2), e635–e635. https://doi.org/10.53761/1.20.02.07

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A Framework for Ethical Integration of Generative AI in Educational Assessment. *Journal of University Teaching and Learning Practice, 21*(06). https://doi.org/10.53761/q3azde36

Ross, M., Bohlmann, J., & Marren, A. (2024). Reflective writing as summative assessment in higher education: A systematic review. *Journal of Perspectives in Applied Academic Practice, 12*(1). https://doi.org/10.56433/jpaap.v12i1.577

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Sharma, S. (2026). Educational integrity in GenAI-augmented assessment: Making judgement visible. *International Journal for Educational Integrity, 22*(1), 6. https://doi.org/10.1007/s40979-026-00216-y

Spies, R., Grobbelaar, S., & Botha, A. (2020). A Scoping Review of the Application of the Task-Technology Fit Theory. In M. Hattingh, M. Matthee, H. Smuts, I. Pappas, Y. K. Dwivedi, & M. Mäntymäki (Eds), *Responsible Design, Implementation and Use of Information and Communication Technology* (Vol. 12066, pp. 397–408). Springer International Publishing. https://doi.org/10.1007/978-3-030-44999-5_33

Stephenson, R., & Armstrong, C. (2026). *Student Generative Artificial Intelligence Survey 2026* (No. 199). HEPI. https://www.hepi.ac.uk/reports/student-generative-ai-survey-2026/

Su, F., O'Dea, M., Wood, M., & Seidl, E. (2026). Embracing a new world: Authentic assessment designs in an age of generative artificial intelligence (GenAI). *Perspectives: Policy and Practice in Higher Education*, 1–12. https://doi.org/10.1080/13603108.2025.2601741

Ucan, S., & Demirel Ucan, A. (2026). Generative AI in Student Writing: A Six–Level Taxonomy for Responsible Use in Higher Education Assessment. *Journal of Academic Ethics, 24(*2), 63. https://doi.org/10.1007/s10805-026-09739-y

Upsher, R., Heard, C., Yalcintas, S., Pearson, J., & Findon, J. (2024). Embracing generative AI in authentic assessment: Challenges, ethics, and opportunities. In *Using Generative AI Effectively in Higher Education*. Routledge.

Wang, Q., Wang, Y., & Amini, M. (2026). Associations of task technology fit and perceived usefulness with responsible generative AI use among university students. *Scientific Reports, 16*(1), 2759. https://doi.org/10.1038/s41598-025-32692-6

Webb-Davies, K., Ratner, S., & Bassett, M. A. (2026). Voice-First Written Assessment (VFWA): Restoring evidence of reasoning through structural redesign (Tyh8z_v1). EdArXiv. https://osf.io/preprints/edarxiv/tyh8z_v1/

Ross, M., Bohlmann, J., & Marren, A. (2024). Reflective writing as summative assessment in higher education: A systematic review. *Journal of Perspectives in Applied Academic Practice, 12*(1). https://doi.org/10.56433/jpaap.v12i1.577

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Sharma, S. (2026). Educational integrity in GenAI-augmented assessment: Making judgement visible. *International Journal for Educational Integrity, 22*(1), 6. https://doi.org/10.1007/s40979-026-00216-y

Spies, R., Grobbelaar, S., & Botha, A. (2020). A Scoping Review of the Application of the Task-Technology Fit Theory. In M. Hattingh, M. Matthee, H. Smuts, I. Pappas, Y. K. Dwivedi, & M. Mäntymäki (Eds), *Responsible Design, Implementation and Use of Information and Communication Technology* (Vol. 12066, pp. 397–408). Springer International Publishing. https://doi.org/10.1007/978-3-030-44999-5_33

Stephenson, R., & Armstrong, C. (2026). *Student Generative Artificial Intelligence Survey 2026* (No. 199). HEPI. https://www.hepi.ac.uk/reports/student-generative-ai-survey-2026/

Su, F., O'Dea, M., Wood, M., & Seidl, E. (2026). Embracing a new world: Authentic assessment designs in an age of generative artificial intelligence (GenAI). *Perspectives: Policy and Practice in Higher Education*, 1–12. https://doi.org/10.1080/13603108.2025.2601741

Ucan, S., & Demirel Ucan, A. (2026). Generative AI in Student Writing: A Six–Level Taxonomy for Responsible Use in Higher Education Assessment. *Journal of Academic Ethics, 24(*2), 63. https://doi.org/10.1007/s10805-026-09739-y

Upsher, R., Heard, C., Yalcintas, S., Pearson, J., & Findon, J. (2024). Embracing generative AI in authentic assessment: Challenges, ethics, and opportunities. In *Using Generative AI Effectively in Higher Education*. Routledge.

Wang, Q., Wang, Y., & Amini, M. (2026). Associations of task technology fit and perceived usefulness with responsible generative AI use among university students. *Scientific Reports, 16*(1), 2759. https://doi.org/10.1038/s41598-025-32692-6

Webb-Davies, K., Ratner, S., & Bassett, M. A. (2026). Voice-First Written Assessment (VFWA): Restoring evidence of reasoning through structural redesign (Tyh8z_v1). EdArXiv. https://osf.io/preprints/edarxiv/tyh8z_v1/

# Artful Early Years: Children's Gendered Awareness and Expression

Michael Brain Martins, Yorkville University, Canada

The European Conference on Education 2026
Official Conference Proceedings

**Abstract**

The gender binary, as a socially constructed framework, continues to shape early childhood experiences in ways that marginalise gender diverse identities. This critical arts-based research (CABR) study examines how children aged 3–9 understood and expressed gender through artistic creation within a community-based setting in Oshawa, Ontario. Grounded in a Pedagogy of Care (POC) and Critical Transformative Learning (CTL), participants created visual representations of gender using loose-parts and engaged in semi-structured interviews and a gallery walk. Qualitative thematic and resonance analysis revealed that these participant children actively negotiate gender through processes of social regulation, imagination, and resistance, rather than passively internalising gender norms. Findings highlight the role of their socialisation agents in reinforcing binary constructs while also demonstrating children's capacity to reimagine gender in fluid and expansive ways. This study contributes to transformative pedagogical approaches that centre relational care, creativity, and inclusive gender possibilities in early years educational settings.

*Keywords:* gender binary, early years children, pedagogy of care, critical transformative learning theory, critical arts-based research, social identity

iafor
The International Academic Forum
www.iafor.org

## Introduction

Transformational equity work in early childhood education requires attention to how gender is produced, regulated, and negotiated through everyday social, linguistic and imaginative practices (Blaise, 2005; Butler, 1990, 2004; MacNaughton, 2000; West & Zimmerman, 1987). Rather than treating gender education as a later corrective intervention, this study examines the micro-politics of belonging, recognition, and power through which participant children encountered gender in relationships, classrooms, media, and play (Ahmed, 2006; B. Davies, 2003; Foucault, 1978; hooks, 1994; Paris & Alim, 2017).

As a gender-fluid educator and researcher, my experience of gendered discrimination and work with Ontario educators through the Principal Qualification Program courses have made me attentive to educators' uncertainty around gender diversity, shaping my commitment to care-based approaches that create opportunities to question, expand, and reimagine gender (Butler, 1990, 2004; B. Davies, 2003; Freire, 1970; Gilligan, 1982; Noddings, 1984, 2002). Early-years education can move beyond performative inclusion when care sustains dialogue, imagination, and belonging happen (Freire, 1970, 1998; Gilligan, 1982; hooks, 1994; Paris & Alim, 2017).

### Situating Me in the Research

Consistent with Creswell and Poth (2018), Lincoln and Guba (1985), and Tracy (2010), positionality is central to qualitative inquiry because researchers bring values, experiences, and assumptions into interpretations. I am genderfluid and use he/him/il pronouns; my Gender Sexual Minority (GSM) identity and experiences within cisnormativity and homophobia heighten my sensitivity to moments in which the gender binary is reinforced or challenged. I also recognise privileges associated with being a White, educated, and able-bodied, male-presenting person, which limit what I know about the broader GSM groups and may shape my interpretations.

I am Canadian-born to immigrant Portuguese parents and raised within a Christian-Judeo cultural context that initially framed gender within the binary. My experiences as an artist, performer, and educator have enabled me to interpret diverse and non-binary representations as meaningful. Drawing on McIntosh's (1998) "invisible backpack," I used reflexive journaling to examine how privilege, emotion, and power operated visibly and invisibly in the research (Foucault, 1978, 1995; hooks, 1994; McIntosh, 1998). Butler's (1990, 2004), Freire (1970), and Kumashiro (2001) further informed my attention to how gender norms become normalised through educational practices.

My preference for gender-fluid possibilities is acknowledged as a bias grounded in an academic and pedagogical commitment to equity, relational justice, and disrupting normative power structures (Blaise, 2005; Butler, 1990, 2004; B. Davies, 2003). This commitment informs my use of Pedagogy of Care (POC), Culturally Relevant/Sustaining Pedagogy (CRP/CSP), and anti-oppressive education towards gender-inclusive practices in early years settings (Freire, 1970, 1998; Gilligan, 1982; Paris & Alim, 2017), Reflexivity and transparency were used not to eliminate bias, but to remain attentive to conformity and resistance and interpret the participant children's meaning-making in their own terms/contexts (Creswell & Poth, 2018; Holmes, 2020; Lincoln & Guba, 1985).

## Theoretical Framework

This study brings together POC and CSP, critical feminist theory (CFT), queer theory (QT), critical transformative learning (CTL), and Foucault's conception of power. Together, these perspectives position gender as relational, socially produced, and open to negotiation while positioning educators as participants in, rather than authorities over, children's gender meaning-making.

### *Pedagogy of Care (POC)*

A POC positions education as relational and ethical, foregrounding attentiveness, connection, responsiveness, and mutual responsibility (Gilligan, 1982; Noddings, 1984, 2002). Davey's (2023) "good enough" teacher extends this orientation by balancing care of students with care for educators. In this study, care creates conditions for dialogue and exploration without immediate correction or categorisation. Freire's (1970) dialogical pedagogy and hooks' (1994) education as a practice of freedom further connect care with social justice in educational practice.

### *Culturally Relevant Pedagogy (CRP)*

CRP/CSP extends culturally relevant and responsive approaches by sustaining and generating educational cultural practice from the ground-up, rather than simply responding to difference(s) (Paris & Alim, 2017). It centers students' voices, lived experience, choice, collaboration, and cultural plurality. In this arts-based gender inquiry, CSP supported open-ended narratives and material choices through participants' children expressed gender using colours, textures, ornamentation, imagery, and storytelling.

### *Critical Feminist and Critical Queer Theories (CFT/QT)*

CFT and QT challenge the binary and heteronormative structures and examine how gender categories become socially intelligible and regulated (Butler, 1990, 2004; hooks, 1994; Jagose, 1996; Sullivan, 2003). Butler's (1990, 2004) account of "performativity" shows how repeated norms reproduce regulation while also creating possibilities for "slippage" and "resignification". Early childhood scholarship similarly demonstrates that children negotiate gender through discourse, play and peer relations (Blaise, 2005; B. Davies, 2003).

### *Critical Transformative Learning (CTL)*

CTL adds reflection and shifts in perspective (Mezirow, 1978, 1991; Taylor, 1997), while critiques of its rational emphasis foreground emotions, symbolism, imagination, and embodied experiences (Boyd, 1989; Boyd & Myers, 1988; Cranton & Roy, 2003; Dirkx, 1998). Accordingly, this study treats transformation as relational and creative, rather than a single cognitive change.

### *Foucauldian Conceptions of Power*

Foucault's (1978, 1995) conception of power provides the framework for regulation. Power is diffuse, productive, and enacted through everyday practices rather than only formal authority. Language, peer judgement, colour coding, play, and social roles can become micro-politics through which gender categories are produced and maintained. Resistance emerges within

relations of power rather than outside them (Foucault, 1978), allowing conformity, imagination, and resistance to be understood as interconnected rather than sequential.

## Research Questions

The main research question guiding this work was: "How do children aged 3–9 understand and express gender identities/constructs through art in their social environments?" Two sub-questions refined this focus:

1. How do 3–9-year-olds interpret gender from interaction with social agents and adults?
2. How can educators unpack and interrupt traditional gender relations among 3–9-year-olds?

The literature review explores these factors and addresses the gaps regarding gender in early years education.

# Literature Review

Gender is widely understood as socially constructed through discourse, practice, interaction, and institutional regulation rather than a fixed binary attribute (Butler, 1990; Crenshaw, 1989; B. Davies, 2003; J. Lorber, 2014, 2018; West & Zimmerman, 1987). Children encounter cultural understandings early through family, peers, educators, and cultural scripts, which can reinforce gendered expectations through everyday interactions (Blaise, 2005; B. Davies, 1989; Leaper & Friedman, 2007; C. L. Martin & D. N. Ruble, 2004). Developmental research has described early gender categorisation as stage-based, with children siding with one of the binaries (Kessler & McKenna, 1978; Kohlberg, 1966; C. L. Martin & D. N. Ruble, 2004; Signorella et al., 1993), while critical early childhood scholarship shows that gender cues and categories are negotiated within social relations rather than simply acquired and fixed (Blaise, 2005; B. Davies, 2003; Halim et al., 2017; Hanish & Fabes, 2014; King et al., 2021; Leaper & Friedman, 2007). The persistence of the binary reflects historical, cultural, and institutional systems that privilege certain identities while marginalising others (Butler, 2004; Collins, 1990; Foucault, 1978; J. Lorber, 2014, 2018).

Educational environments can reproduce gender expectations through language, materials, grouping, and play. Educators may unintentionally encourage nurturing behaviours for girls and competitive or physical behaviours for boys (Balter et al., 2018; A. Davies et al., 2021; MacNaughton, 2000; B. Martin, 2011). Socialisation can occur through subtle feedback rather than explicit teaching, shaping what children perceive as recognisable or acceptable gender behaviour (Butler, 1990, 2004; Cherney et al., 2003; Foucault, 1978, 1995; Freeman, 2007; hooks, 1994; B. Martin, 2011). Media is another significant socialising agent, with historical and recent research identifying stereotypical representations alongside some marketable-diverse moments (Bailey, 2009; Chapman, 2021; Clarke et al., 2024; Davey, 2022; Giroux, 1999; Picker, 2002; Rollero et al., 2023; Santoniccolo et al., 2023; Skjelstad & Ellefsen, 2024; Terán & Conroy, 2025). Children therefore encounter gender messages across interconnected social environments, not only educational settings (Blaise, 2005; B. Davies, 2003; MacNaughton, 2000; Muchabaiwa & Makonye, 2025; 2000; Renold, 2005; West & Zimmerman, 1987).

## Gaps in the Research

Despite increasing attention to gender-inclusive education, practical research on everyday early-years implementation remains limited (Balter et al., 2018, 2021; A. Davies et al., 2021; K. H. Robinson & C. Davies, 2017). More research is needed on how contemporary media, social media, digital platforms, and artmaking influence gender meaning-making in early childhood (Buckingham, 2013; King et al., 2021; Picker, 2002; Santoniccolo et al., 2023; Scarcelli et al., 2021), and how educators balance gender inclusivity with social and cultural resistance (Barone & Eisner, 2012; Beren, 2013; Trout et al., 2022). This study addresses these gaps through participant children's art, language, social relationships, and interaction, with attention to pedagogical applicability.

# Methodology

I employed critical arts-based research (CABR), integrating critical theory with participatory and creative practices to examine gender meaning-making with participant children (Janmohamed, 2010; Leavy, 2023a, 2023b). CABR supports multimodal, participatory inquiry in which art was treated as a way of knowing rather than simply an illustration of verbal responses (Barone & Eisner, 2012; Eisner, 1965; Leavy, 2023a, 2023b; Springgay et al., 2005). This design aligned with the Ontario Kindergarten Curriculum (OKC), care, co-creation, multimodality, and ethical engagement through collage with loose-parts, semi-structured interviews, and a gallery walk (Davey, 2023; Ontario Ministry of Education, 2016; Trout et al., 2019, 2022).

### *Research Design*

The three primary methods were (a) collage with loose-parts for artmaking, (b) semi-structured artwork-based interviews, and (c) final gallery walks. The design encouraged participant children to construct meaning of a girl, a boy, and a person through open-ended materials and minimal preset prompting, allowing visual and verbal meaning-making to develop together (Leavy, 2023a, 2023b).

### *Participants and Context*

Six children (four females, two males), aged 3–9, participated in the study within a community-based rental space in Oshawa, Ontario. Recruitment expanded from the original age range because of recruitment constraints. Participants represented diverse backgrounds of age, gender, race, and socioeconomic status, and engaged voluntarily with parental consent and ongoing daily assent through the four research sessions (Table 1). While all participants initially identified within the gender binary, their expressions revealed more nuanced understandings of gender throughout the study.

**Table 1**
*Participant Demographic Characteristics*

| Participants Number | Gender Disclosed During Research | Age | Race |
|---|---|---|---|
| Participant 1 (P1) | female | 3 years old | non-racialised |
| Participant 2 (P2) | female | 5 years old | non-racialised |
| Participant 3 (P3) | female | 8 to 9 years old | non-racialised |
| Participant 4 (P4) | female | 4 years old | racialised |
| Participant 5 (P5) | male | 9 years old | mixed-race |
| Participant 6 (P6) | male | 3 years old | mixed-race |

*Note.* An overview of recruited participant demographics that situate their multiple identities.

### *Methods and Data Collection*

Data were created across four sessions. In the first three sessions, participant children created loose-parts collage art representing a girl, a boy, and any person they wished, followed by an individual interview. The fourth session was a collective gallery walk where all participant art pieces were displayed, and participants discussed their artwork with peers and the counsellor (adult proxy), while I observed and recorded field notes as a non-participant observer. Each session included renewed assent, reminders of support, and opportunities for breaks (V. Braun & V. Clarke, 2006; V. Braun & V. Clarke, 2021; Cahnmann-Taylor & Siegesmund, 2018; Miles & Huberman, 1994).

### *Data Analysis*

Data analysis combined visual and thematic approaches through qualitative research resonance (Creswell & Poth, 2018; Good, 2023; Tracy, 2010). I examined patterns in transcripts and field notes alongside visual, verbal and non-verbal features (weak codes), including colours, textures, spatial placement, ornamentation, layering, imagery, pauses, gestures, and narrative explanations (Saldaña, 2015) (Figure 1).

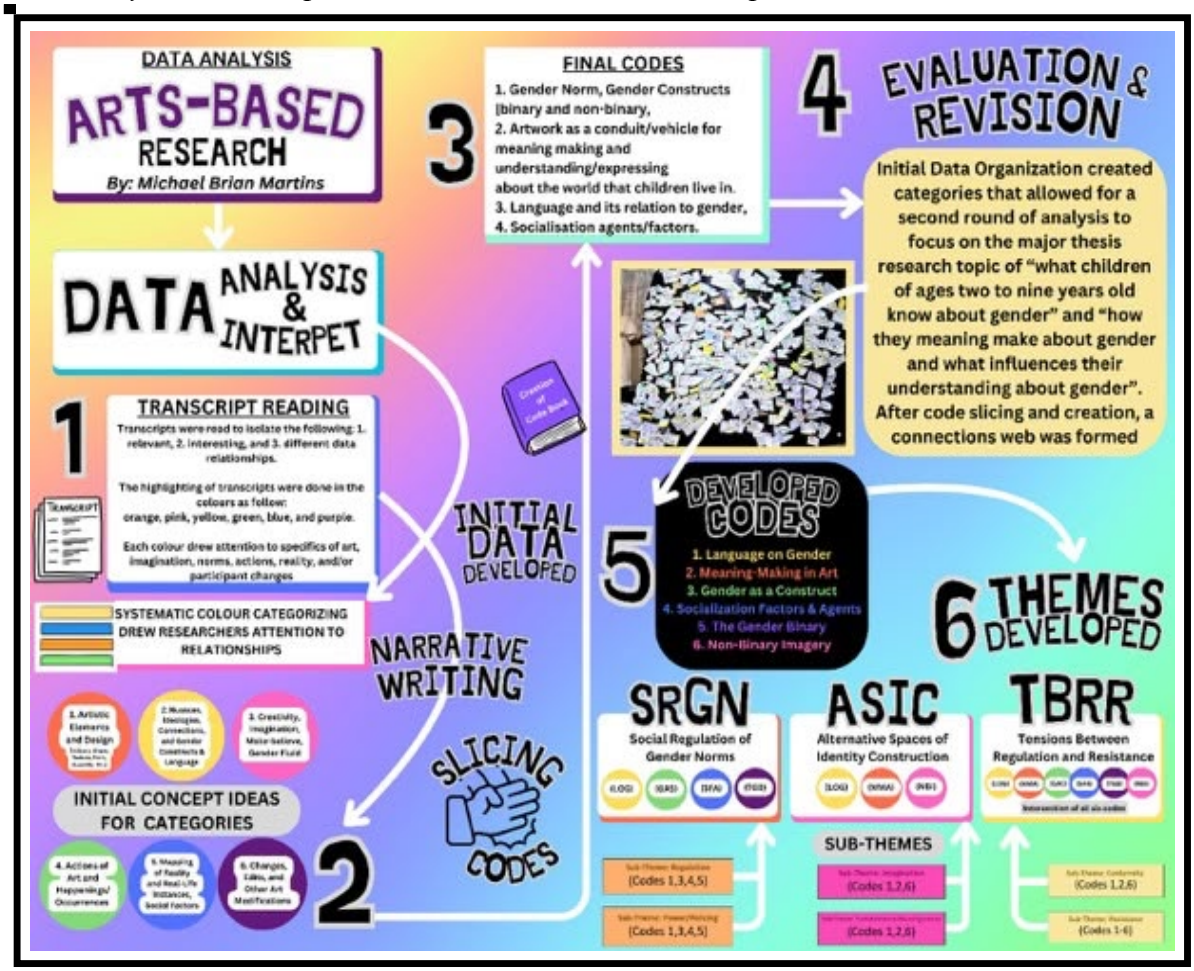
**Figure 1**

*Code Slicing and Mapping During Data Analysis and Categorisation (Full Board–Overview)*

*Note.* This figure visualises the categorisation process, using colored paper to show weak and strong code connections and coloured yarn to outline how these codes converge into themes.

Open coding incorporated inductive and deductive processes, followed by iterative clustering into six stronger codes: Meaning-Making in Art (MMA), Language on Gender (LOG), Socialisation Factors and Agents (SFA), Gender as a Construct (GAC), The Gender Binary (TGB), and Non-Binary Imagery (NBI) (Fife & Gossner, 2024; Saldaña, 2015). Attention was given to shifts in meaning across contexts, highlighting the relational nature of gender construction. Constant comparison across artwork, interviews, and the gallery walk observation supported interpretive and thematic coherence (V. Braun & V. Clarke, 2006; V. Braun & V. Clarke, 2021; Fife & Gossner, 2024; Miles & Huberman, 1994; Miles et al., 2013; Saldaña, 2015) (Figure 2).

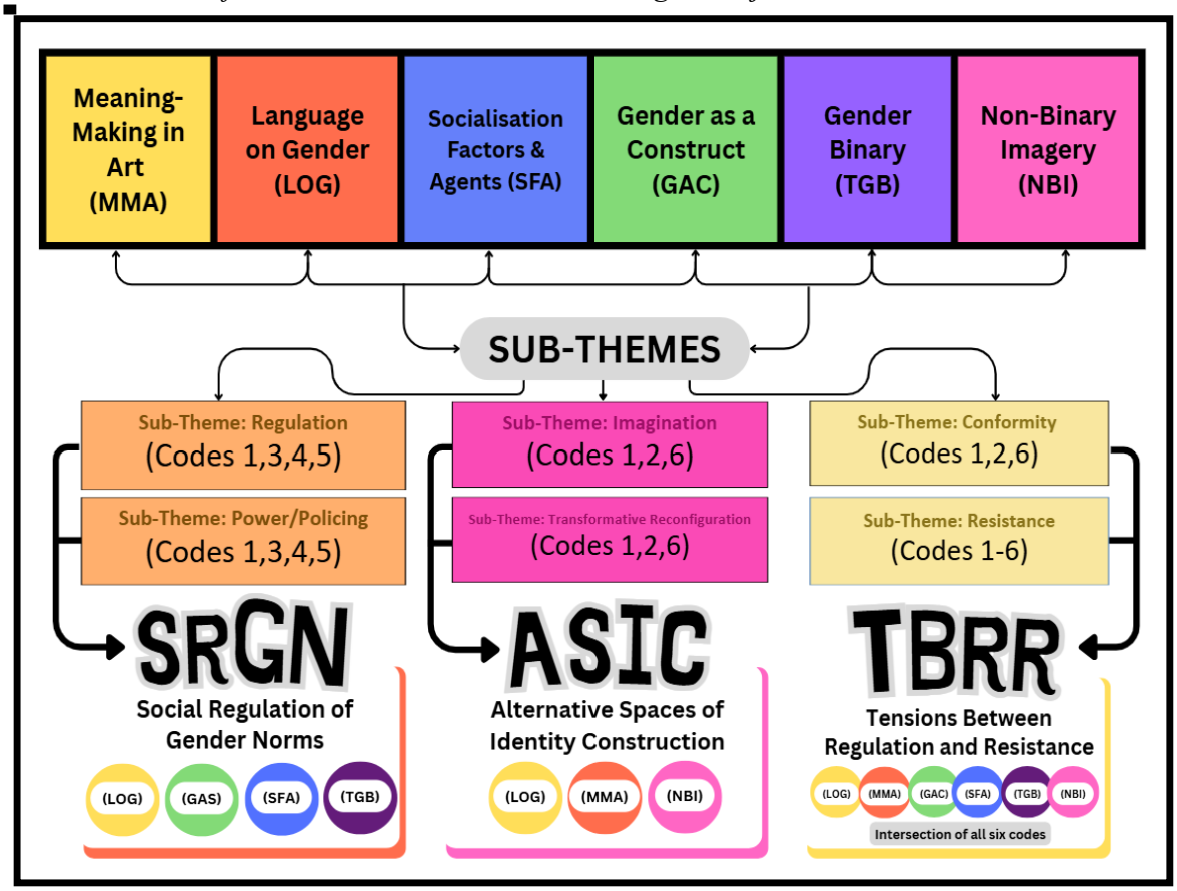
**Figure 2**
*Data Analysis and Interpretation: Code Formation and Explanation*




*Note.* This figure outlines the data analysis process, illustrating the procedural flow, interconnection, and interpretation of codes as they culminate in sub-themes and major themes.

The six stronger codes were connected through 29 weaker codes, creating six sub-themes: regulation, power/policing, imagination, transformative reconfigurations (TR), conformity and resistance. These developed into three major themes: Social Regulation of Gender Norms (SRGN), Alternative Spaces of Identity Construction (ASIC), and Tensions Between Regulation and Resistance (TBRR). Visual and verbal data were treated as interconnected because artistic features could carry meanings that language alone might not access (Figure 3).

**Figure 3**
*Interconnection of Codes to Minor Themes, Leading to Major Themes*

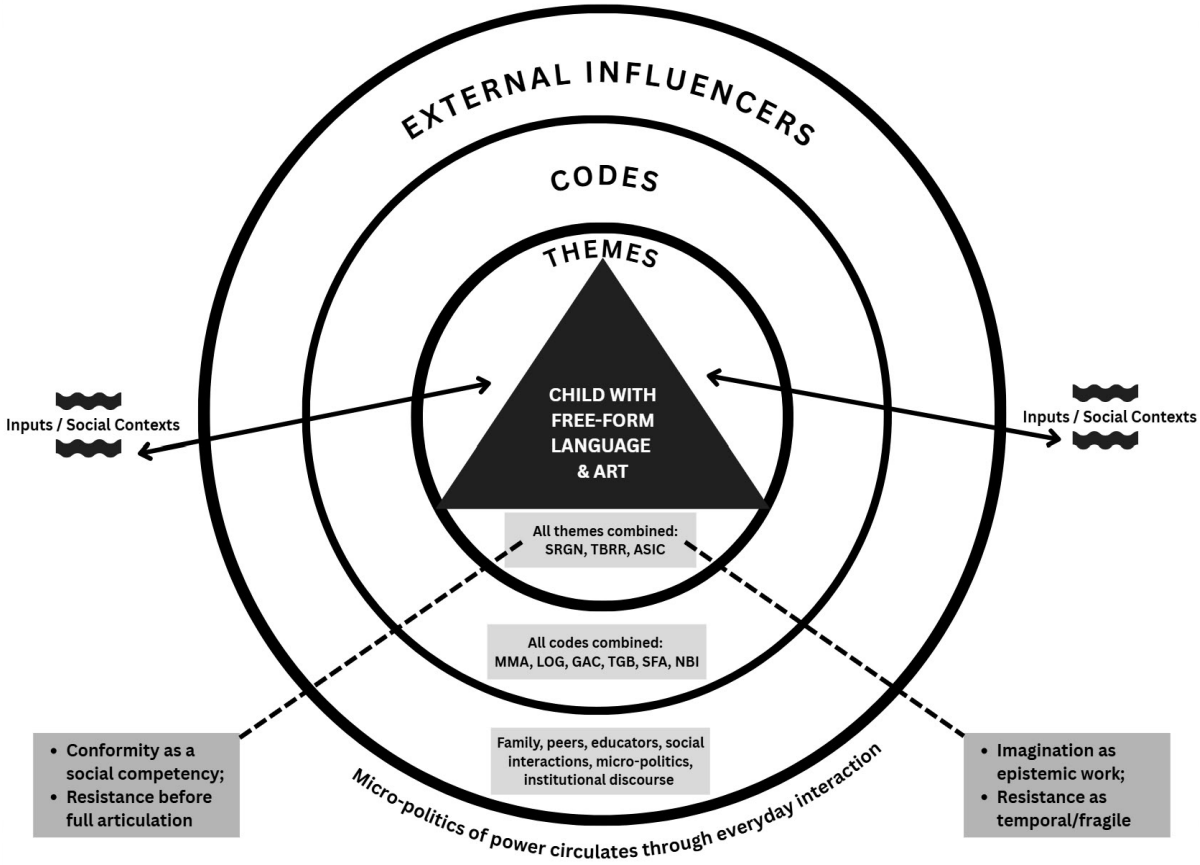



*Note.* This figure demonstrates the progression of codes leading to sub-themes and how these codes and minor themes work in conjunction, forming the thesis's major themes.

## Conceptual Framework (TF)

This study's conceptual framework (CF), akin to the atom/molecule analogy, conceptualises gender as relational, dynamic, and embedded within power. Figure 4 centres participant children's language, communication, imagination, and artmaking potential; surrounding layers represent the codes, sub-themes, socialisation agents, and broader social contexts. The three themes: SRGN, ASIC, and TBRR interact fluidly in a cumulative, non-linear manner; regulation may prompt imagination, imagination may create resistance, and resistance may return to conformity or generate further reconfigurations. This layered approach makes visible how art creations, play, narratives, and interactions inform, with credibility, the study's conceptual claims.

**Figure 4**
*Conceptual Framework: Regulation, Imagination, Resistance (Networked/Bowl Model)*




## Ethical Considerations

Ethical care included parental consent, renewed child assent, confidentiality, voluntary participation, withdrawal option, secure data storage, REB approval, and counsellor support. The counsellor remained present to support risk management, breaks and debriefing, while open-ended interview prompts and children-driven artmaking helped minimise leading or harmful interactions. Amendments to the age range and sessions were approved by REB before implementation. Participant information was anonymised, and parents could request transcripts as needed. Personal-reflexive journaling and professional support managed researcher positionality and emotions, recognising care as an ethical responsibility to participants and the researcher.

## Clarification on Conflation of Biological Sex and Gender

Throughout the findings, these children often used “girl” and “boy” as linguistic placeholders for culturally learned notions of femininity and masculinity, rather than as scholarly distinctions between biological sex and gender identity (Butler, 1990, 2004; B. Davies, 1989, 2003; C. L. Martin & D. N. Ruble, 2004; Signorella et al., 1993). This language conflation reflects developmental and socially produced binary patterns and should therefore be understood as participants’ everyday categories, and not essentialist understandings (Grossman, 2023; Joyce, 2021; Sax, 2026; Shrier, 2020; Stock, 2021).

## Findings

The findings revealed that these children's gender meaning-making was neither fixed nor uniformly binary. Across the data, participants reproduced familiar gender markers while also using art, language, imagination, and interactions to expand and renegotiate gender actively, beyond developmental arguments of age appropriateness (Blaise, 2005, 2009; Cherney et al., 2003; MacNaughton, 2000; National Association for the Education of Young Children [NAEYC], 2020; Paris & Alim, 2017; Thorne, 1993). The analysis developed six interconnected code clusters—MMA, LOG, SFA, GAC, TGB, and NBI—which informed the three major themes below.

### *Theme 1: Social Regulation of Gender Norms (SRGN)*

Participant children frequently used familiar gender markers in artwork, including colour, hair, body size, clothing, ornamentation, family roles, and language. The MMA, LOG, GAC, and TGB codes showed how visual and verbal signs worked together to position characters within familiar feminine or masculine categories. Participant One used bows and lipstick to signal femininity, while Participant One linked their father being "bigger than mommy," and Participant Three said that beards, body hair, and larger size were masculine traits. Furthermore, Participant Four stated, "strong and fights bad guys." These artistic attributes categorised traits into gender expectations. Participant Five explained gender role catergorisation in their "Minion girl" character was "waiting for her husband" (Figure 5). These examples demonstrate patterns in how gender categories become intelligible through socially recognisable artistic and linguistic cues (Butler, 1990, 2004; B. Davies, 2003; MacNaughton, 2000).

**Figure 5**
*Children's Gendered Feminine and Masculine Art Pieces*

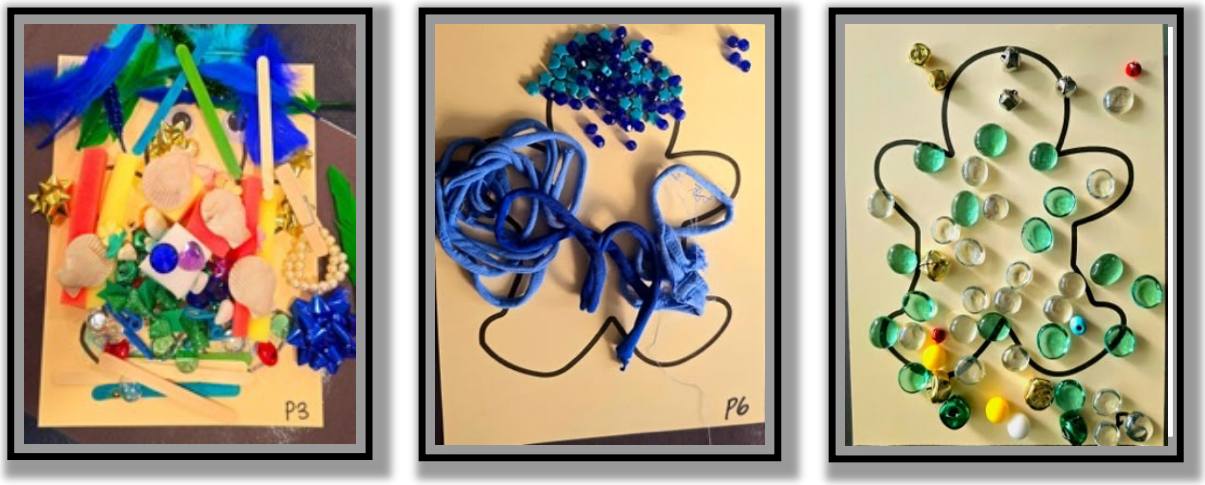
These distinctions were often presented as natural or self-evident, suggesting that repeated interactions with social agents normalised gender expectations for these children (Butler, 1990, 2004; B. Davies, 2003). Participant Five's artwork connected language, art, and socialisation and the gender binary, demonstrating how objects carried learned relational meaning. Together, these examples illustrated SRGN as everyday regulation rather than explicit instruction in the absorption of gender norms (S. Nicholson, 1971).

### *Theme 2: Alternative Spaces of Identity Construction (ASIC)*

Alongside regulation, artmaking created spaces for imaginative experimentation. Participant children used rainbow colours, crystals, shine, magical characterisation, super-heroism, and hybrid forms to move beyond familiar gender markers. As Barone & Eisner (2012), Eisner (1965), and Springgay et al. (2005) suggested, art can function as inquiry cognitively, where participant artistic choice allowed them to assign and reassign meaning through materials and visual symbols, not always inherently gendered.

Participant Five explained, “The shine means you’re a different kind of boy,” transforming a familiar masculine colour marker into a more flexible identity. Participants Three and Five described pink and blue as available to any gender (influenced by peer/parents). Participants One and Two used rainbow colours for characters they described as “a girl, a boy, or a person.” Participant Six’s Sky Giant and other fantastical figures further demonstrated that imaginative worlds could suspend ordinary gender rules by blending “Supersonic” and “Glenda” into the giant with both genders. In doing so, many participants used terms such as “beautiful” and “handsome” in ways that crossed conventional gendered language (Figure 6).

**Figure 6**
*Mixed Participant Artwork Exhibiting Alternative Constructions*




These moments resonate with J. Nicholson and K. Steele’s (2020) emphasis on resistant imagination and Tsai’s (2012) account of creative play. As participant children navigated their imagination alongside learned gender notions, social contexts actively shaped their new learning (Hickey-Moody, 2013; Renold, 2005; Ringrose, 2013; Skeggs, 1997, Walkerdine, 1990, 1997). Participant children shifted between binary and non-binary expressions, demonstrating that gender was continually negotitated and that imagination could function as an epistemic resource with new possibilities (Butler, 1990, 2004; Greene, 1995; Leavy, 2023a, 2023b).

These children’s responses indicate that when given the opportunity, they used binary norms merely as a starting point to understand and explore genderfluidity. Understood as an epistemic practice(s), art creations did not simply reflect internal beliefs but reworked them through social interactions with others and participants' social world Butler, 1990, 2004; Renold, 2005; Ringrose, 2013; Skeggs, 1997; Walkerdine, 1990, 1997). For example, Participant Three stating, “boy hair and girls' clothes,” “can be for both,” and Participant One simply stating their

art as “just a person.” They knew this because they said, “all people can wear clothes and rainbow colours”, displaying deliberate resistance to gender categorisation (Figure 7). Participant Four hybridised names in fantastical imagined concepts, blending “Supersonic and Glenda” to map/mix “amazing qualities,” creating “Sky Giant” akin to their dad. The findings indicate that participants were actively testing, combining, and reconfiguring gender in regulated social contexts with agency (Greene, 1995; Leavy, 2023b).

**Figure 7**
*Children’s Flexible Gender Art Pieces*

***Theme 3: Tensions Between Regulation and Resistance (TBRR)***

Resistance was often fragile and rarely appeared as outright rejection of gender norms. Participant Three stated, “Not all boys have short hair,” and recalled, “Once my dad wore a dress,” disrupting a gendered marker without rejecting the category of boy. Participant Five retained that “blue means you’re a boy” while adding that shine meant a “different kind of boy” and blended “beautiful” and “handsome” across genders, fluidly understanding body and ornamentation markers. Participant Six reworked “shine” and “sparkle” as attributes available beyond femininity, applying them to the masculine (Figure 6). These moments illustrate partial resistance through reconfiguration using “slippage” moments for new gender possibilities (Butler, 1990, 2004; Foucault, 1978).

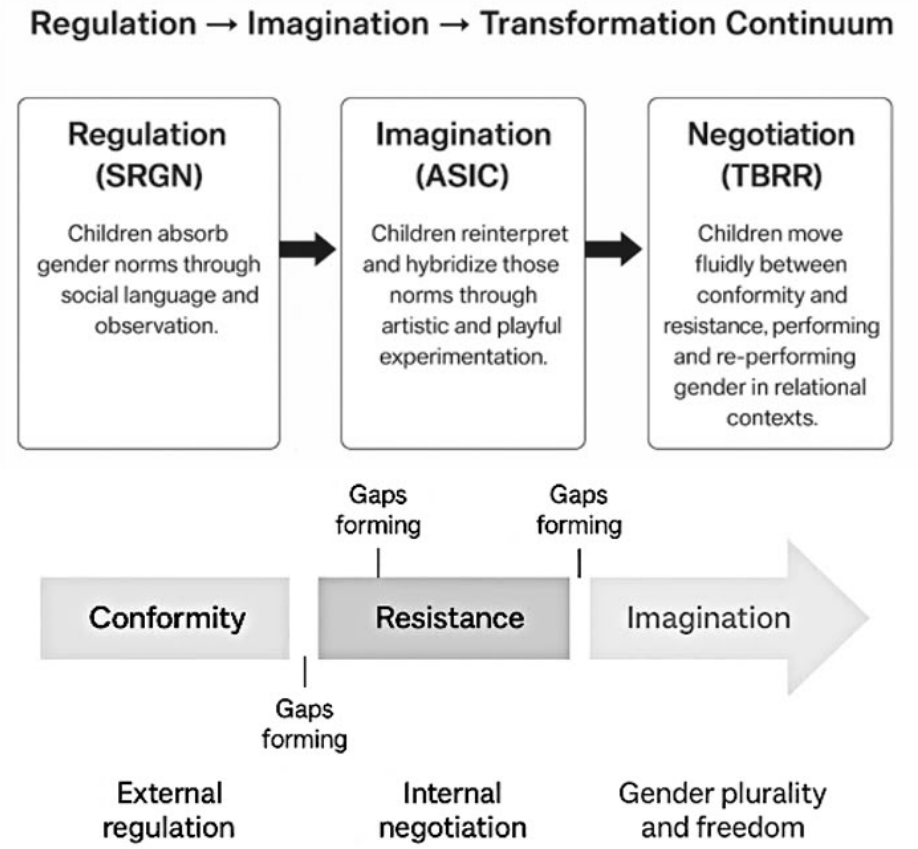
The gallery walk demonstrated how relational contexts could shift meaning. Non-judgmental interactions created openings for participant children to explain, reconsider, and extend their gender representations, including when Participant Three connected Pride, rainbow, and mixed gender markers as safe zones of resistance to binary constraints through hybrid, fantastical imagery, where various colours, sparkles, rainbows, bows, and lipstick can be for all (Greene, 1995; Jagose, 1996; Sullivan, 2003) (Figure 6 & 7). These moments suggest that resistance can emerge through an individual's flexible gender position using dialogue and social context rather than a fixed, stage-based conformist position (Greene, 1995; Foucault, 1978; hooks, 1994; Kohlberg, 1966; C. L. Martin & D. N. Ruble, 2004). Often, the binary functioned as a strategic mechanism for social intelligibility rather than a developmental limitation.

These findings suggest that resistance is an emergent, relational process shaped by complex contexts and micro-politics of social interaction (Foucault, 1978). “Fragile resistance,” drawing on Blaise (2005), B. Davies (2003), and MacNaughton (2000), captures moments when several

participants reworked gender meanings without fully rejecting familiar categories. Resistance therefore appeared as a possibility that could strengthen, recede, or re-emerge across participant interactions, where a non-linear continuum across themes SRGN, ASIC, and TBRR opens new gender possibilities (Figure 8).

**Figure 8**

*Continuum of Early Years Understanding About Gender Through Themes and Concepts*




*Note*. Regulation (SRGN) structures gender learning; imagination (ASIC) offers openings—*cracks*—within the gender structure: resistance (TBRR) transforms openings into possibilities for participants' new gender understandings.

## Discussion

The findings indicate that early-years gender meaning-making is socially regulated, relationally negotiated, and imaginatively mediated by participant children. Gender appeared less as an individual cognitive milestone than as something learned through repeated interactions with socialisation agents, supporting West and Zimmerman's (1987) "doing gender", Butler's (1990, 2004) accounts of performativity, and early childhood scholarship outlining how children negotiate gender within social relations, and not only developmentally (Blaise, 2005; B. Davies, 2003; R. W. Connell, 1995; R. W. Connell, 2000; Kessler & McKenna, 1978; Kohlberg, 1966; C. L. Martin & D. N. Ruble, 2004; Muchabaiwa & Makonye, 2025; Signorella et al., 1993; Thorne, 1993). Art simultaneously reiterated and destablised gender norms, showing how power could operate both regulatorily and productively (Butler, 1990; Foucault, 1978; hooks, 1994).

First, social regulation operated through ordinary practices rather than only explicit instruction: colour, bodily descriptors, relational roles, labels, and peer judgements became gendered cues

in the micro-politics of gender (Foucault, 1978). Bourdieu's (1984) habitus explains how repeated classifications of gender can be made intelligible by participants (Bhana, 2007; Blaise, 2005, 2009; Hanish & Fabes, 2014; MacNaughton, 2000; Thorne, 1993). The regulatory politics of belonging were interactional by policing gender while recategorizing it (Bali, 2021; Cherney et al., 2003; Freeman, 2007; Grant & Pittaway, 2022; B. Martin, 2011). Second, conformity should not automatically be interpreted as passive compliance; participants sometimes used familiar gender categories to secure recognition, social competence, and belonging within the binary system (Butler, 1990, 2004; Clarke et al., 2024; R. W. Connell, 1995; R. W. Connell, 2000; Skeggs, 1997; Schiele et al., 2020; Terán & Conroy, 2025; Walkerdine, 1990, 1997).

This study examines the binary opposition between biological essentialism and social constructionism. While essentialist critiques view gender as fixed or adult-induced (Grossman, 2023; Joyce, 2021; Shrier, 2020; Stock, 2021), these findings show participants actively reworking social scripts. This aligns with NAEYC's (2020) updated Developmentally Appropriate Practice statement, which prioritises context-dependent, sociocultural interpretations over universal age stages (Terán & Conroy, 2025). Consequently, resistance manifested as child-centred agency through repetition, slippage, and reconfiguration. Conformity and resistance fluctuated across contexts, while artmaking and children's dialogues created relaxed openings for imaginative reconfigurations (Barone & Eisner, 2012; Eisner, 1965; Greene, 1995; Leaper & Friedman, 2007; Ruble et al., 2006; Springgay et al., 2005). Art provided fluid pathways for new learning where verbal language was constrained or socially risky (Atkinson & DePalma, 2009; Butler, 1990, 2004; Tsai, 2012). Through hybrid representations, these children constructed inclusive visual metaphors for fluidity, creating "safe zones" where binary rules dissolved (Greene, 1995; Jagose, 1996; Sullivan, 2003).

Third, imagination functioned as an epistemic resource (Leavy, 2023a, 2023b). Participants' use of colours, textures, ornamentation, hybrid imagery, and fantastical narratives to express understandings not always fully articulated by speech provides new gender communication (Barone & Eisner, 2012; Eisner, 1965; Greene, Springgay et al., 2005). In communicating about gender, these children did not first "know" gender and then illustrate it; rather, they simultaneously absorbed social knowledge, externalised, and reworked the knowing through the material act of creation (Hickey-Moody, 2013; Leavy, 2023b; J. Nicholson & K. Steele, 2020). Fourth, resistance emerged as fragile and situated: participants reworked familiar categories artistically rather than rejecting the binary entirely. "Blue means you're a boy" could coexist with a "different kind of boy"; "not all boys" have short hair; and "beautiful" and "handsome" could be blended across genders (Barker & Scheele, 2016; Blaise, 2005). These moments support Foucauldian and Butlerian accounts of resistance and resignification while challenging claims that gender flexibility is necessarily adult-induced child-compliance (Grossman, 2023; Joyce, 2021; Shrier, 2020; Stock, 2021).

For educators, the implication is not to correct children's understandings from above but to recognise where regulation, conformity, imagination, and resistance are already occurring. Practical responses include using "everyone" rather than "boys and girls," organising groups without gender categories, keeping play materials open-ended, and responding to gender policing with calm, shame-free questions such as, "Why do you think that?" or "Could it be for anyone?" These practices preserve children's agency while widening the social possibilities grounded in POC, education as a practice of freedom, and CRP/CSP (Freire, 1970; Gilligan, 1982; hooks, 1994; Paris & Alim, 2017).

Affirming environments also require gender-inclusive literature, unlabelled play areas, and loose-parts, while remaining attentive to institutional and political constraints in learning spaces (Ahmed, 2006; Windle & Rosa, 2023). Because participants drew on familiar media characters and cultural scripts, educators can invite children to redesign characters, change roles, or imagine different relationships, using questions such as, "Could the superhero be caring?" or "Could the princess save herself?" (Greene, 1995). Teaching that is open-ended and context-responsive supports hybridisation and avoids closures (Gay, 2010; Ladson-Billings, 1995; Paris, 2012; Paris & Alim, 2017). Finally, these findings support critical media literacy.

### *Conceptual Framework and Continuum Synthesis*

This conceptual framework brings the findings together as a network rather than a sequence (Figure 4 & 8). SRGN represents the social pressures and scripts; ASIC represents the imaginative spaces through which meaning can expand; and TBRR represents the movement between conformity and resistance. Participant children did not move predictably from the binary understanding to fluidity. They could conform in one moment, imagine differently in another, and resist or reconfigure a gender norm later. The continuum is therefore recursive: regulation may prompt imagination; imagination may produce resistance; resistance may recede into conformity; and conformity may provide recognisability from which further reconfigurations become possible for children in their social contexts.

## Limitations

This exploratory qualitative study has several limitations. The sample was small, context-specific, and drawn from one regional community, so the findings are not statistically generalisable. None of the participants explicitly identified outside the gender binary, limiting exploration of a wider range of gender identities. The open-ended arts-based design produced varying engagement, especially among younger participants, and interpretation of visual data involved the researcher's judgement. Family perspectives were largely absent, only examined through participant narratives. The study was also not longitudinal, did not systematically examine cognitive processes, and did not analyse digital media as a distinct social agent. These limitations point to future comparative, longitudinal, family-engaged, and media-focused research.

## Significance and Implications

This study contributes to early childhood gender research by integrating social regulation, imagination, and resistance within an interpretive framework grounded in child-centred arts-based data aligned with OKC inclusion practices (Ontario Ministry of Education, 2016). It challenges deficit interpretations of children's binary gender talk by showing that conformity and flexibility can coexist and demonstrates how artmaking can reveal affective, relational, and conceptual meanings that language alone may not access fully (Barone & Eisner, 2012; Leavy, 2023a, 2023b; Springgay et al., 2005).

Pedagogically, early childhood learning environments are active sites of gender learning. Language, play, materials, character creation, and adult or peer interactions can reinforce or disrupt gender norms. A care-centred, culturally relevant/sustaining approach therefore requires more than symbolic representation; it requires conditions in which children can explore difference, question assumptions, and participate in meaning-making. Inclusion is strongest when representation is embedded daily rather than presented as a one-off activity,

along with educator preparation and institutional support to respond to gender policing without avoiding the gender topic (Balter et al., 2018, 2021; A. Davies et al., 2021).

### Recommendations

Five recommendations emerged. First, engaging in critical gender pedagogy earlier: children are already interpreting gender through art, play, language, media and relationships, so gender education should not be deferred until adolescence. Second, ground gender education in POC and cultural sustainability by centring children's lived experiences, dialogue, choices, belonging, and relational safety (Davey, 2023; Gilligan, 1982; Paris & Alim, 2017). Third, design spaces that interrupt regulation through open-ended art, loose-parts, diverse materials, and inclusive language. Fourth, integrate critical media literacy by inviting children to remix familiar characters and question gendered assumptions (Atkinson & DePalma, 2009; Greene, 1995; Muchabaiwa & Makonye, 2025; NAEYC, 2020; S. Nicholson, 1971; Terán & Conroy, 2025; Tsai, 2012). Fifth, strengthen teacher education and institutional advocacy so educators can translate critical theory into practical action and respond to gender policing without shame or avoidance (Balter et al., 2018; 2021; A. Davies et al., 2021).

## Conclusion

This study demonstrates that early-years children’s gender meaning-making is socially regulated, relationally negotiated, and imaginatively mediated. Across SRGN, ASIC, and TBRR, participants produced familiar norms using art, language, imagination, and interaction to renegotiate gender boundaries. Conformity supported social intelligibility, imagination functioned as an epistemic resource, and resistance emerged as partially situated, not outright rejection.

Educators can support more inclusive gender learning by recognising these processes rather than treating children’s expressions as errors to correct. Grounded in POC, CRP/CSP, critical inquiry, and arts-based practice, educators can sustain the imaginative openings through which children can explore identity. Gender diversity in early years is not an exceptional topic, but part of a complex social early childhood experience.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, I (the author) used AI-assistive technology to identify text redundancies and to improve the language, stylistic clarity, and readability of this paper. Following that process, I refined the content to ensure accuracy and remain fully responsible for the originality, conceptual framework, and academic integrity of the final paper submission.

## References

Ahmed, S. (2006). Orientations: Toward a queer phenomenology. *GLQ: A Journal of Lesbian and Gay Studies*, *12*(4), 543–574. https://static1.squarespace.com/static/58ad660603596eec00ce71a3/t/58bec800b8a79b7c599de24a/1488898050432/Orientations+Toward+a+Queer+Phenomenology.pdf

Atkinson, E., & DePalma, R. (2009). Un-believing the matrix: Queering consensual heteronormativity. *Gender and Education*, *21*(1), 17–29. https://www.tandfonline.com/doi/abs/10.1080/09540250802213149

Bailey, C. (2009). Embracing the icon: The feminist potential of the trans bodhisattva, Kuan Yin. *Hypatia*, *24*(3), 178–196. https://doi.org/10.1111/j.1527-2001.2009.01051.x

Bali, M. (2021, April 10). Pedagogy of care – Caring for teachers. Reflecting Allowed. https://blog.mahabali.me/pedagogy/critical-pedagogy/pedagogy-of-care-caring-for-teachers/

Balter, A. S., Van Rhijn, T., & Gores, D. (2021). Supporting the development of sexuality in early childhood: The rationales and barriers to sexuality education in early learning settings. *The Canadian Journal of Human Sexuality, 30*(3). https://utppublishing.com/doi/abs/10.3138/cjhs.2021-0034

Balter, A. S., Van Rhijn. T., & Davies, A. W. J. (2018). Equipping early childhood educators to support the development of sexuality in childhood: Identification of pre-and post-service training needs. *The Canadian Journal of Human Sexuality*, *27*(1), 33–42. https://utppublishing.com/doi/abs/10.3138/cjhs.2017-0036?role=tab

Barker, M. J., & Scheele, J. (2016). *Queer: A graphic history*. Icon Books.

Barone, T., & Eisner, E. W. (2012). *Arts based research*. Sage.

Beren, M. (2013). Gay and lesbian families in the early childhood classroom: Evaluation of an online professional development course. *LEARNing Landscapes*, *7*, 61–79. https://doi.org/10.36510/learnland.v7i1.630

Bhana, D. (2007). The price of innocence: Teachers, gender, childhood sexuality, HIV and AIDS in early schooling. *International Journal of Inclusive Education, 11*(4), 431–444. https://doi.org/10.1080/13603110701391394

Blaise, M. (2005). *Playing it straight: Uncovering gender discourses in the early childhood classroom*. Routledge.

Blaise, M. (2009). “What a girl wants, what a girl needs”: Responding to sex, gender, and sexuality in the early childhood classroom. *Journal of Research in Childhood Education*, *23*(4), 450–460. https://doi.org/10.1080/02568540909594673

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste* (R. Nice, Trans.). Harvard University Press.

Boyd, R. D. (1989). Facilitating personal transformations in small groups Part I. *Small Group Behavior, 20*(4), 459–474. https://doi.org/10.1177/104649648902000406

Boyd, R. D., & Myers, J. G. (1988). Transformative education. *International Journal of Lifelong Education*, *7*(4), 261–284. https://doi.org/10.1080/0260137880070403

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE.

Buckingham, D. (2013). *Media education: Literacy, learning and contemporary culture.* Polity Press.

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge.

Butler, J. (2004). *Undoing gender*. Routledge. https://selforganizedseminar.wordpress.com/wp-content/uploads/2011/07/butler-undoing_gender.pdf

Cahnmann-Taylor, M., & Siegesmund, R. (2018). *Arts-based research in education. Foundations for practice.* Routledge.

Chapman, R. (2021). Moving beyond 'gender-neutral': Creating gender expansive environments in early childhood education. *Gender and Education*, *34*(1), 1–16. https://doi.org/10.1080/09540253.2021.1902485

Cherney, I. D., Kelly-Vance, L., Gill Glover, K., Ruane, A., & Oliver Ryalls, B. (2003). The effects of stereotyped toys and gender on play assessment in children aged 18–47 months. *Educational Psychology*, *23*(1), 95–106. https://doi.org/10.1080/01443410303222

Clarke, L. L., Hine, B., England, D., Flew, P. P. M. S., Alzahri, R., Juriansz, S. N., & Garcia, M. J. B. C. (2024). The gendered behaviors displayed by Disney protagonists. *Frontiers in Sociology*, *9*, Article 1338900. https://doi.org/10.3389/fsoc.2024.1338900

Collins, P. H. (1990). Black feminist thought in the matrix of domination. In P. H. Collins (Ed.), *Black feminist thought: Knowledge, consciousness, and politics of empowerment* (pp. 221–238). Unwin Hyman.

Connell, R. W. (1995). *Masculinities*. University of California Press.

Connell, R. W. (2000). *The men and the boys*. University of California Press.

Cranton, P., & Roy, M. (2003). When the bottom falls out of the bucket: Toward a holistic perspective on transformative learning. *Journal of Transformative Education, 1*(2), 86–98. https://doi.org/10.1177/1541344603001002002

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist policies. *University of Chicago Legal Forum, 1*, 139–167. https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1052&context=uclf

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry & research design. Choosing among five approaches* (4th ed.). Sage.

Davey, N. (2022, February 15). How watching cartoons saved my family—and my sanity. *Chatelaine.* https://chatelaine.com/living/entertainment/cartoons-saved-my-family-and-my-sanity/

Davey, N. (2023). *A pedagogy of care.* Brill. https://doi.org/10.1163/9789004547520_004

Davies, A. W. J., Simone-Balter, A., & van Rhijn, T. (2021). Sexuality education and early childhood educators in Ontario, Canada: A Foucauldian exploration of constraints and possibilities. *Contemporary Issues in Early Childhood*, *24*(4), 394–410. https://doi.org/10.1177/14639491211060787

Davies, B. (1989). *Frogs and snails and feminist tales: Preschool children and gender*. Hampton Press.

Davies, B. (2003). *Shards of glass: Children reading and writing beyond gendered identities*. Hampton Press.

Dirkx, J. M. (1998). Transformative learning theory in the practice of adult education: An overview. *PAACE Journal of Lifelong Learning*, *7*, 1–14.

Eisner, E. W. (1965). Curriculum ideas in a time of crisis. *Art Education*, *18*(7), 7–12. https://doi.org/10.2307/3190712

Fife, S. T., & Gossner, J. D. (2024). Deductive qualitative analysis: Evaluating, expanding, and refining theory. *International Journal of Qualitative Methods*, *23*, 1–12. https://doi.org/10.1177/16094069241244856

Foucault, M. (1978). *The history of sexuality: Vol. 1. An introduction* (R. Hurley, Trans.). Pantheon Books. (Original work published 1976)

Foucault, M. (1995). *Discipline and punish: The birth of the prison* (A. Sheridan, Trans.). Vintage Books.

Freeman, N. K. (2007). Preschoolers' perceptions of gender appropriate toys and their parents' beliefs about genderized behaviors: Miscommunication, mixed messages, or hidden truths? *Early Childhood Education Journal*, *34*(5), 357–366. https://doi.org/10.1007/s10643-006-0123-x

Freire, P. (1970). *Pedagogy of the oppressed*. Herder & Herder.

Freire, P. (1998). *Pedagogy of freedom: Ethics, democracy, and civic courage*. Rowman & Littlefield.

Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice*. Teachers College Press.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Harvard University Press.

Giroux, H. A. (1999). *The mouse that roared: Disney and the end of innocence*. Rowman & Littlefield.

Good, H. (2023, February 12). A guide to coding qualitative research data [Web Page]. *Dovetail Editorial Team*. https://dovetail.com/research/qualitative-research-coding/

Grant, A., & Pittaway, S. (2022). *Enacting a pedagogy of kindness*. Routledge.

Greene, M. (1995). *Releasing the imagination: Essays on education, the arts, and social change*. Jossey-Bass.

Grossman, M. (2023). *Lost in Trans-nation: A child psychiatrist's guide out of the madness*. Skyhorse.

Halim, M. L. D., Ruble, D. N., Tamis-LeMonda, C. S., Shrout, P. E., & Amodio, D. M. (2017). Gender attitudes in early childhood: Behavioral consequences and cognitive antecedents. *Child Development*, *88*(3), 882–899. https://doi.org/10.1111/cdev.12642

Hanish, L. D., & Fabes, R. A. (2014). Peer socialization of gender in young boys and girls. In R. E. Tremblay, M. Boivin, & R. DeV Peters (Eds.), *Encyclopedia on early childhood development* [online]. https://www.child-encyclopedia.com/gender-early-socialization/according-experts/peer-socialization-gender-young-boys-and-girls

Hickey-Moody, A. (2013). *Youth, arts, and education: Reassembling subjectivity through affect*. Routledge. https://doi.org/10.4324/9780203855829

Holmes, A. G. D. (2020). Researcher positionality—A consideration of its influence and place in qualitative research—A new researcher guide. *Shanlax International Journal of Education*, *8*(4), 1–10. https://doi.org/10.34293/education.v8i4.3232

Hooks, B. (1994). *Teaching to transgress: Education as the practice of freedom*. Routledge.

Jagose, A. (1996). *Queer theory: An introduction*. New York University Press.

Janmohamed, Z. (2010). Queering early childhood studies: Challenging the discourse of developmentally appropriate practice. *Alberta Journal of Educational Research*, *56*(3), 304–318. https://doi.org/10.11575/ajer.v56i3.55413

Joyce, H. (2021). *Trans: When ideology meets reality*. OneWorld.

Kessler, S. J., & McKenna, W. (1978). *Gender: An ethnomethodological approach*. University of Chicago Press.

King, T. L., Scovelle, A. J., Meehl, A., Milner, A. J., & Priest, N. (2021). Gender stereotypes and biases in early childhood: A systematic review. *Australasian Journal of Early Childhood, 46*(2), 112–125. https://doi.org/10.1177/1836939121999849

Kohlberg, L. (1966). *A cognitive-developmental analysis of children's sex-role concepts and attitudes.* Center for Moral Education.

Kumashiro, K. K. (2001). Perspectives on anti-oppressive education in social studies, English, mathematics, and science classrooms. *Educational Researcher*, *30*(3), 3–12. https://doi.org/10.3102/0013189X030003003

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, *32*(3), 465–491. https://doi.org/10.3102/00028312032003465

Leaper, C., & Friedman, C. K. (2007). The socialization of gender. In J. E. Grusec, & P. D. Hastings (Eds.), *Handbook of socialization: Theory and research* (pp. 561–587). The Guilford Press.

Leavy, P. (2023a). *Re/Invention: Methods of arts-based research* (2nd ed.). Guilford Press.

Leavy, P. (2023b). *Research design: Quantitative, qualitative, mixed methods, arts-based, and community-based participatory research approaches* (2nd ed.). The Guilford Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Lorber, J. (1994). *Paradoxes of gender*. Yale University Press.

Lorber, J. (2018). Constructing gender: The dancer and the dance. In J. A. Holstein & J. F. Gubrium (Eds.), *Handbook of constructionist research* (pp. 1–34). Guilford Press. https://doi.org/10.4324/9780429499821-61

MacNaughton, G. (2000). *Rethinking gender in early childhood education*. Sage.

Martin, B. (2011). *Children at play: Learning gender in the early years.* Trentham Books.

Martin, C. L., & Ruble, D. N. (2004). Children's search for gender cues: Cognitive perspectives on gender development. *Current Directions in Psychological Science*, *13*(2), 67–70. https://doi.org/10.1111/j.0963-7214.2004.00276.x

McIntosh, P. (1998). White privilege: Unpacking the invisible knapsack. In M. McGoldrick (Ed.), *Re-visioning family therapy: Race, culture, and gender in clinical practice* (pp. 147–152). The Guilford Press. (Reprinted from "Peace and Freedom," July/August 1989)

Mezirow, J. (1978). Perspective transformation. *Adult Education Quarterly*, *28*, 100–110. https://doi.org/10.1177/074171367802800202

Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass.

Miles, M. B., & Huberman, M. A. (1994). *Qualitative data analysis* (2nd ed.). Sage.

Miles, M. B., Huberman, M. A., & Saldaña, J. (2013). *Qualitative data analysis: A methods sourcebook*. Sage.

Muchabaiwa, W., & Makonye, L. (2025). The influence of digital media on gender stereotypes in ECD learners: Leveraging heavy media consumption by young children to demystify gender stereotypes. *International Journal of Research and Innovation in Social Science*, *9*(3SEDU), 6515–6526. https://doi.org/10.47772/IJRISS.2025.903SEDU0476

National Association for the Education of Young Children [NAEYC]. (2020). *Developmentally Appropriate Practice: A Position Statement of the National Association for the Education of Young Children*. NAEYC. https://www.naeyc.org/resources/position-statements/dap/contents

Nicholson, J., & Steele, K. (2020). *Radically listening to transgender and gender expansive young children: Creating epistemic justice through critical reflection and resistant imaginations* [Monograph]. In Libra. https://doi.org/10.5771/9781498590389

Nicholson, S. (1971). How not to cheat children: The theory of loose parts. *Landscape Architecture*, *62*(1), 30–34.

Noddings, N. (1984). *Caring, a feminine approach to ethics & moral education*. University of California Press.

Noddings, N. (2002). *Starting at home: Caring and social policy*. University of California Press.

Ontario Ministry of Education. (2016). *The Kindergarten program*. Government of Ontario. https://www.ontario.ca/document/kindergarten-program-2016

Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational Researcher, 41*, 93–97. https://doi.org/10.3102/0013189X12441244

Paris, D., & Alim, H. S. (2017). *Culturally sustaining pedagogies*. Teachers College Press.

Picker, R. (Director). (2002). *Mickey Mouse monopoly* [Video]. Media Education Foundation. https://www.kanopy.com/centennialcollege/video/mickey-mouse-monopoly-0?frontend=kui

Renold, E. (2005). *Girls, boys and junior sexualities: Exploring children's gender and sexual relations in the primary school*. Routledge Falmer.

Ringrose, J. (2013). *Postfeminist education? Girls and the sexual politics of schooling*. Routledge.

Robinson, K. H., & Davies, C. (2017). Sexuality education in early childhood. In L. Allen, & M. Rasmussen (Eds.), *The Palgrave handbook of sexuality education* (pp. 217–242). Palgrave Macmillan.

Rollero, C., Teresi, M., & Pagliaro, S. (2023). The role of sexting on the perception of image-based sexual abuse. *Journal of Interpersonal Violence*, *38*(21–22), 11727–11744. https://doi.org/10.1177/08862605231188131

Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (2006). Gender development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (6th ed., pp. 858–932). John Wiley & Sons.

Saldaña, J. (2015). *The coding manual for qualitative researchers*. Sage.

Santoniccolo, F., Trombetta, T., Paradiso, M. N., & Rollè, L. (2023). Gender and media representations: A review of the literature on gender stereotypes, objectification and sexualization. *International Journal of Environmental Research and Public Health*, *20*(10), Article 5770. https://doi.org/10.3390/ijerph20105770

Sax, L. (2026). *Why gender matters: What parents and teachers need to know about the emerging science of sex differences* (2nd ed.). Basic Books.

Scarcelli, C., Krijnen, T., & Nixon, P. (2021). Sexuality, gender, media. Identity articulations in the contemporary media landscape. *Information, Communication & Society*, *24*(8), 1063–1072. https://doi.org/10.1080/1369118X.2020.1804603

Schiele, K., Louie, L., & Chen, S. (2020). Marketing feminism in youth media: A study of Disney and Pixar animation. *Business Horizons*, *63*(5), 659–669. https://doi.org/10.1016/j.bushor.2020.05.001

Shrier, A. (2020). *Irreversible damage: The transgender craze seducing our daughters*. Regnery.

Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children's gender schemata about others: A meta-analytic review. *Developmental Review, 13*(2), 147–183. https://doi.org/10.1006/drev.1993.1007

Skeggs, B. (1997). *Formations of class and gender: Becoming respectable*. Sage.

Skjelstad, E., & Ellefsen, L. W. (2024). Challenging stereotypes? Norwegian music teachers' repertoires on gender roles and gender-expansiveness. *Journal of Research in Music Education*, *71*(4), 398–417. https://doi.org/10.1177/00224294231175859

Springgay, S., Irwin, R. L., & Kind, S. W. (2005). A/r/tography as living inquiry through art and text. *Qualitative Inquiry*, *11*(6), 897–912. https://doi.org/10.1177/1077800405280696

Stock, K. (2021). *Material girls: Why reality matters for feminism*. Fleet.

Sullivan, N. (2003). *A critical introduction to queer theory*. New York University Press.

Taylor, E. W. (1997). Building upon the theoretical debate: A critical review of the empirical studies of Mezirow's transformative learning theory. *Adult Education Quarterly, 48*(1), 34–59. https://doi.org/10.1177/074171369704800104

Terán, L., & Conroy, M. (2025). *GDI's 2025 children's TV report: Record number of female leads in new shows*. Geena Davis Institute. https://geenadavisinstitute.org/research/gdis-2025-childrens-tv-report-record-number-of-female-leads-in-new-shows/

Thorne, B. (1993). *Gender play: Girls and boys in school*. Rutgers University Press.

Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative Inquiry, 16*(10), 837–851. https://doi.org/10.1177/1077800410383121

Trout, İ. Y, Pennington, B., & Christensen, M. C. (2019). Connecting classroom to community through photovoice: Pedagogical implications. *LEARNing Landscapes*, *12*(1), 285–301. https://doi.org/10.36510/learnland.v12i1.992

Trout, L., Wilks, J., Holkham, L., & Reid, C. (2022). Arts-based research methods in educational research: Lessons from practice. *International Journal of Research & Method in Education, 45*(2), 134–149.

Tsai, K. C. (2012). Play, imagination, and creativity: A brief literature review. *Journal of Education and Learning, 1*(2), 15–20. http://dx.doi.org/10.5539/jel.v1n2p15

Walkerdine, V. (1990). *Schoolgirl fictions*. Verso.

Walkerdine, V. (1997). *Daddy's girl: Young girls and popular culture*. Harvard University Press.

West, C., & Zimmerman, D. H. (1987). Doing gender. *Gender and Society, 1*(2), 125–151. https://doi.org/10.2307/189945

Windle, J., & Rosa, J. (2023). Navigating critical language education at the crossroads of neonationalism and neoliberalism: The role of border talk. *TESOL Quarterly, 57*(3). https://doi.org/10.1002/tesq.3229

**Contact email:** michaelbmartins@yahoo.ca

© The International Academic Forum 2026
The International Academic Forum (IAFOR)
Sakae 1-16-26-201
Naka Ward, Nagoya, Aichi
Japan 460-0008
www.iafor.org