Transforming Teaching and Learning at the Royal University of Bhutan Through Targeted Needs Assessment

Gembo Tshering, Royal University of Bhutan, Bhutan

The Asian Conference on Education 2024
Official Conference Proceedings

Abstract
To ensure ongoing relevance and effectiveness, teaching and learning in higher education institutions must undergo regular reviews and targeted interventions. This study undertook a comprehensive needs assessment to guide the development of targeted professional development (PD) programs in teaching, learning, and assessment in the nine colleges of the Royal University of Bhutan. Employing a convergent parallel mixed-methods design, data were collected from 318 faculty members through a mixed-methods survey questionnaire. This instrument integrated established tools, including the Approaches to Classroom Assessment Inventory Version 3, Teaching Competencies Scale, and Revised Approaches to Teaching Inventory, ensuring a robust evaluation of faculty needs. Typically, needs assessment studies use one of these instruments, resulting in a partial understanding of the needs. The findings revealed a significant and pressing demand for PD initiatives focusing on student-centered teaching strategies, innovative assessment approaches, using assessment data to inform instructional activities, providing qualitative feedback, and understanding 21st-century learners. Furthermore, the study uncovered considerable variation in professional development needs across the different colleges, underscoring the necessity for customized and tailored PD programs to each institution's unique contexts and challenges, highlighting the critical importance of continuous, need-based professional development initiatives. The study also discusses the implications of these findings for designing and implementing effective PD programs, offering insightful and practical recommendations for institutional leaders and policymakers. Drawing upon a framework built from this study, PD initiatives at the Royal University of Bhutan are now being implemented and funded by EUFSTIAT, a project funded by the Erasmus Plus Programme.

Keywords: Professional Development, Assessment Techniques, Royal University of Bhutan, Classroom Assessment Inventory

Introduction

Context and Importance of Professional Development

In the current educational landscape, characterized by rapid technological advancements and shifting pedagogical paradigms, professional development (PD) is crucial for educators (Darling-Hammond et al., 2017; Jacob et al., 2015). PD programs help faculty members stay updated with the latest teaching methodologies, learning technologies, and assessment strategies, ultimately leading to improved educational outcomes (Bergmark, 2020; Sims & Flecher-Wood, 2020). Despite the availability of general PD programs, there is a growing recognition of the need for tailored PD that addresses specific institutional and disciplinary needs (Kohan et al., 2023; Steinert, 2016; van Dijk, 2023).

Purpose of Needs Assessment

The primary purpose of this study was to conduct a needs assessment across nine colleges, using existing tools from the literature, to identify the specific gaps in teaching, learning, and assessment. The outcomes of this assessment are intended to guide the development of customized PD programs that directly address the identified needs, ensuring that faculty members receive the support necessary to enhance their instructional practices and student engagement.

Research Gap

While the literature well-documents the importance of PD (Darling-Hammond et al., 2017; Jacob et al., 2015), there is a paucity of research on needs assessments that focus specifically on teaching, learning, and evaluation across multiple institutions (Kohan et al., 2023; Steinert, 2016; van Dijk, 2023). This study seeks to fill this gap by providing empirical data on faculty needs across nine distinct colleges, contributing to the broader understanding of how PD programs can be better designed and implemented.

Literature Review

Theoretical Framework

This study is grounded in the theory of andragogy, as articulated by Knowles and colleagues (2015), whose seminal work has been cited over 27,000 times. The theory outlines six fundamental principles, emphasizing that as adults mature, they become increasingly independent and self-directed, taking charge of their learning decisions. Adults leverage their extensive life experiences to comprehend new concepts and skills more effectively, making these experiences a crucial asset in learning. They are incredibly motivated to learn when the content directly relates to their personal or professional lives, addressing immediate needs or challenges. With a problem-centered orientation, adults focus on acquiring practical skills that can be immediately applied rather than purely theoretical knowledge.

Additionally, adults are primarily driven by internal motivators, such as the desire for self-improvement and personal growth. For educators, applying andragogy in teaching requires fostering active participation, utilizing learners' experiences, and ensuring the content's relevance and immediate applicability. In this context, needs assessment is essential for identifying these specific learning needs and ensuring that professional development (PD)

programs are relevant and effective (Behar-Hosenstein et al., 2014; Malicka et al., 2017). This study aligns with pedagogical principles by focusing on teaching, learning, and assessment, promoting PD tailored to faculty members' unique needs.

Previous Studies

Research has consistently shown that effective PD programs are essential for improving teaching, learning, and assessment practices of the faculties of higher education institutions and enhancing student learning outcomes. For instance, a study by Darling-Hammond and colleagues (2017) found that effective PD programs are ongoing, focused on specific content areas, and aligned with educators' needs. Similarly, Desimone (2009) and Malicka and colleagues (2017) highlight the importance of PD, which is directly linked to instructional practices and student learning. Despite these findings, many PD programs still need to be more generic and connected to the actual needs of faculty, underscoring the importance of conducting needs assessments (Othayman et al., 2022).

Challenges and Criticisms

Despite the recognized importance of needs assessments, several challenges exist in their implementation (Othayman et al., 2022). One significant challenge is ensuring that the data collected is accurate and reflective of the actual needs of the faculty (Othayman et al., 2022). This requires using validated instruments and careful data analysis (Bastos et al., 2014). Additionally, there is often a disconnect between the findings of needs assessments and the actual implementation of PD programs, with some institutions failing to translate identified needs into actionable PD initiatives because of the disconnect (Parry-Jones & Soulsby, 2002). Parry-Jones and Soulsby (2002) state that the disconnect arises due to an unclear concept of needs and a need for a straightforward assessment framework. This study used three different data collection tools to determine the needs of the Royal University of Bhutan faculty members to counter the risk, as cited by Parry-Jones and Soulsby (2002).

Methodology

Research Design

The needs assessment used a convergent parallel mixed-methods approach (Creswell, 2021). The data were collected using mixed-methods survey questionnaires (Creswell & Hirose, 2019), providing a comprehensive understanding of faculty needs. The study utilized existing validated instruments to ensure reliability and validity in data collection.

Data Collection

Data were collected from 318 faculty members across nine colleges. The survey instruments included sections on teaching practices, learning strategies, and assessment beliefs. These surveys were administered in person, ensuring broad participation. The questionnaire also contained qualitative survey items intended to enable the participants to write their perspectives, which were not included in the quantitative survey items.

The study used the Approaches to Classroom Assessment Inventory Version 3 (ACAI V3) (CART, 2019). The ACAI-V3 instrument is divided into three parts, with the current study focusing on Part C, which addresses assessment beliefs (CART, 2019). Part C is structured

into four primary dimensions: assessment purpose, assessment process, assessment fairness, and assessment theory. Each dimension comprises three specific priority areas. The purpose of assessment encompasses the assessment of learning, assessment for learning, and assessment as learning. The assessment process involves designing, using/scoring, and communicating assessments. Assessment fairness covers standards, equity, and differentiation in assessment practices. Assessment theory pertains to consistency, contextualization, and balance within the theoretical framework of assessment. In addition to the four primary dimensions, Part C incorporates an additional dimension labeled "assessment beliefs."

Specific questionnaire items are developed for each priority area, culminating in 32 items within Part C (CART, 2019). The distribution of items is as follows: assessment of learning, assessment for learning, and assessment as learning each has two questions; design, use/scoring, and communication each has two questions; standards, equity, and differentiation each has two questions; and consistency, contextualization, and balance each has two questions. The "assessment beliefs" dimension is more extensively covered, with eight dedicated questions. Each has seven levels: strongly agree, agree, disagree, strongly disagree, and do not know.

The surveys were distributed through a direct, in-person approach, ensuring thorough engagement with the participants. A dedicated team of researchers visited each of the nine colleges to administer the paper-and-pencil questionnaires. Before distributing the surveys, the researchers conducted a brief orientation session to inform participants about the survey's purpose, objectives, and significance. This session was essential for clarifying any questions and securing informed participation. Participants were provided consent forms after this briefing, which they were asked to review and complete. After collecting the consent forms, the researchers distributed the survey questionnaires to the faculty members. This method allowed for immediate clarification of doubts and ensured a high response rate. Participation was robust, with all faculty members, except those on long-term study leave, participating in the survey. This hands-on approach facilitated a comprehensive data collection process, minimizing the potential for non-response bias and enhancing the reliability of the findings. The response rate was 52%.

Sample

Data were collected using a census approach, ensuring comprehensive participation across the faculty. An initial invitation letter was sent to all faculty members, encouraging their involvement in the survey. Faculty members on campus during the survey administration participated directly in the study. For those off-campus at the time, arrangements were made to facilitate their participation upon their return. Paper-and-pencil questionnaires were left at their respective colleges, allowing these faculty members to complete the survey at their convenience. This method ensured that the study captured a diverse sample, representing faculty from various disciplines, including the humanities, social sciences, and STEM fields, thereby enhancing the generalizability of the findings across different academic areas.

The survey was conducted across nine colleges affiliated with the Royal University of Bhutan, covering a range of academic disciplines. The participating institutions were Paro College of Education (46 respondents), Gedu College of Business Studies (37), College of Science and Technology (35), Samtse College of Education (31), Jigme Namgyel College of Engineering (44), Sherubtse College (33), Gyalpozhing College of Information Technology

(14), College of Language and Cultural Studies (32), and the College of Natural Resources (45). In total, 317 out of 610 faculty members participated in the survey. The diverse fields, including education, business, engineering, information technology, and natural resources, provided a comprehensive overview of the university's academic environment. Table 1 shows the demographic variables of the participants.

Table 1: Demographic Variables

Demographic variables	Frequency %
Gender	
Male	71.8
Female	28.2
Age	
Below 25	4.8
25-29	17.4
30-39	29.7
40-49	27.4
50-59	19.7
Above 60	1.00
Academic Degree	
Bachelor	20.5
Master	63.5
PhD	16.5
Others	4.4
Pfofessional Degree	
Do not have	39.1
Bed (Primary)	1.9
BEd (Secondary)	4.1
PGCE	44.5
MEd	10.4
Teaching Experience	
One Year or Less	9.7
Two Years	7.1
Three Years	6.5
Four years	2.6
Five years	3.2
More than five years	71.0

Data Analysis

Data collected using the Approaches to Classroom Inventory Assessment Version Three (ACAI-V3) were analyzed to compare and synthesize insights into faculty members' beliefs about assessment.

Alignment of Constructs.

The study aimed to determine whether significant differences exist among the faculty members of the nine colleges' beliefs about assessments. The following steps were undertaken in SPSS version 29 to analyze the data:

- The dataset consisted of responses from 317 participants across nine colleges. Each participant's responses to the 32 items were coded on a scale of 1 to 6, with do not coded 7, with higher scores indicating stronger agreement.
- The 'Do not Know' response option was treated as missing data after it was confirmed to be missing completely at random (MCAR), as indicated by Little's MCAR test, $\chi^2(2244)=2326.520$, $p=.110$.

Factor Analysis.

Factor analysis was conducted to identify the underlying structure of the 32 items designed to measure assessment beliefs across multiple dimensions. The goal was to reduce the data into a smaller set of factors that explain the observed correlations among the variables.

Tests for homogeneity of variance were conducted to assess the consistency of variance across different dimensions. A p-value greater than 0.05 typically indicates homogeneity. In this analysis, only three dimensions showed p-values below 0.05, as detailed in Table 2.

Table 2: Dimensions With Levene Statistics Showing P-Values Below 0.05

Dimension	Based on Mean	Levene Statistics	df1	df2	p-value
Assessment Fairness	Based on Mean	3.046	8	8	.003
Assessment Theory	Based on Mean	2.372	8	8	.017
Assessment Purppse	Based on Mean	2.611	8	8	.009

The Kolmogorov-Smirnov and Shapiro-Wilk tests were conducted to assess the normality of the factors across the nine colleges. Generally, a p-value greater than 0.05 suggests that the data do not significantly deviate from normality. In this analysis, most colleges exhibited p-values greater than 0.05, indicating normality. However, the following dimensions and colleges showed significant deviations from normality, as detailed in Table 3.

Table 3: Kolmogrov-Smirno and Shapiro-Wilk Tests

Dimension	College	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistics	df	p	Statistics	df	p
Assessment Process	Sherubtse College	.254	20	.001	.873	20	.013
Assessment Fairness	College of Language and Cultural Studies	.268	19	<.001	.813	19	.002
Assessment Theory	College of Science and Technology	.313	15	<.001	.869	15	.033
	Sherubtse College	.206	20	.026	.874	20	.014
Assessment Purpose	Samtse College of Education	.212	11	.180	.799	11	.009

The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.79, indicating that the sample was suitable for factor analysis. Bartlett's Test of Sphericity was significant ($\chi^2(496)=1828.17$, $p<.001$), confirming that the correlations among variables were sufficient for the analysis.

Principal Component Analysis was employed to extract factors from the 32 items. An initial analysis was run to obtain eigenvalues for each factor in the data. Nine factors had

eigenvalues greater than one and explained 64.23% of the variance. Based on the scree plot, nine factors were retained for further analysis.

Nine components were extracted based on their eigenvalues (greater than 1), explaining 64.218% of the variance. The initial unrotated solution showed that the first component had an eigenvalue of 6.720, accounting for 20.999% of the variance. The second component had an eigenvalue of 3.803, explaining an additional 11.884% of the variance, with the first two components cumulatively explaining 32.883%. Subsequent components explained progressively less variance, with the ninth component accounting for 3.215%, resulting in a cumulative variance of 64.218% across the nine components.

A Varimax rotation was applied to improve the factor structure's interpretability. After rotation, the variance explained by the first component was reduced to 11.999%, and the second component explained 11.177% of the variance. The rotated solution redistributed the variance evenly across the components, resulting in a more apparent factor structure. The first nine components collectively explained 64.218% of the total variance after rotation, with the variance contributions of the components now more balanced.

The principal component analysis (PCA) was conducted to explore the underlying factor structure of the 32 items related to faculty's assessment beliefs. The analysis employed Varimax rotation to enhance the interpretability of the components. The results of the rotated component matrix are presented in Table 4.

Table 4: Rotated Factor Loadings

Item	Component								
	1	2	3	4	5	6	7	8	9
1	.783			.227					
2	.717			.199		-.181		.228	.180
3	.697	.174			.135		.183		.156
4	.647			.181		-.117	.185	.111	
5	.567	.515			-.122	.198			.114
6	.551	.420			.263	.191	.118		
7	.546							.346	.502
8	.488	.186		.317					
9		.836						.100	
10	.124	.675			.141		.152		
11	-.116	.669	.172	.176		.210		.340	
12	.137	.642	.174	.149	.251				.150
13		.532		.522				.249	.137
14		.482	.188	.202	.163	.239	.217	-.138	-.356
15			.826			.173			
16		.135	.776		-.249			.205	
17			.717			.108	.107	-.218	.212
18	.230			.711				.122	
19	.326	.126		.548			.166	-.174	.364
20	.270	.188	.214	.547	.110	-.136			
21	.249	.175	-.242	.469	-.258	.386			
22		.139			.839	.150			-.123
23		.388	.126		.614	-.146	-.118	.165	.353
24	.251	.180	-.234		.595		.385		
25		.182	.236			.794			
26	-.207	.135	.528	-.127		.576		.257	.103
27	-.150		.195	.202	.290	.494		.185	.403
28			.116				.833		.126

Table 4 (cont.)

29	.230		.333	-.136	.633	.217	
30	.223	.140	.167	.115	.178	.670	-.186
31	.194	.289	.261			.607	.217
32	.103	.239			.235	-.112	.662

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 11 iterations.

Nine distinct components were identified through the analysis. Upon close examination of the ACAI V3 items and its literature, Component 1 is identified as Assessment Purpose, Component 2 as Assessment Process, Component 3 as Assessment Fairness, Component 4 as Assessment Theory 1, Component 5 as Assessment Theory 2, Component 6 as Negative Perceptions of Assessment, Component 7 as Time and Overuse of Assessment, Component 8 as Motivation and Positive Impact of Assessment, and Component 9 as Assessment as an Enjoyable Experience. However, it is essential to note that these actual dimensions, as revealed through exploratory analysis, differ from the expected dimensions reported in the ACAI V3 model. Table 5 shows the items and components.

Table 5: Components and Items

Components	Item No.	Items
Assessment Purpose	1	The primary purpose of classroom assessment is to assign a grade or level to student work.
	2	Classroom assessment should be used to determine if students have met programme standards.
	3	Feedback from classroom assessments improve student learning.
	4	By using assessment, teachers can track the progress of students.
	5	Students should use assessments to evaluate their own work.
	6	Students are able to provide accurate and useful feedback to each other.
	7	Teachers have the skills and knowledge to construct good assessments.
	8	For good classroom assessment, teachers need extensive knowledge of the curriculum.
Assessment Process	9	Classroom assessment involves judging a student's performance in relation to a set of goals/standards/criteria.
	10	Classroom assessment is integral to developing lesson plans and implementing curriculum.
	11	Classroom assessment should be used to provide evidence of student progress for administrative purposes.
	12	Classroom assessment is useful when reporting a student's achievement/progress to parents and caregivers.
	13	An important component of classroom assessment is students taking largescale tests (e.g., provincial assessment, EQAO).
	14	Provincial assessments (e.g., EQAO) are a meaningful form of assessment.
Assessment Fairness	15	All classroom assessments should be adapted to suit the learning needs of identified students (e.g., English language learners).
	16	Students with exceptionalities should be provided with different classroom assessment than other students.
	17	For good assessment, teachers need to know how their students learn.

Table 5 (cont.)

Assessment Theory 1	18	Classroom assessment helps teachers identify the particular learning needs of any student.
	19	Classroom assessment results provide reliable information.
	20	Classroom assessment involves teachers making judgements about how well a student is learning in relation to other students.
	21	Classroom assessment results is a good indicator of the quality of a school.
Assessment Theory 2	22	Assessment results reflect the quality of teaching.
	23	Assessment is an imprecise process.
	24	Observing students is a valid form of assessment.
Negative Perceptions of Assessment	25	Classroom assessment is of little use to teachers on a day-to-day basis.
	26	Classroom assessment interrupts students' learning.
	27	Assessment is a stressful activity for students.
Time and Overuse of Assessment	28	Assessment takes time away from teaching.
	29	Teachers use too many assessments.
Motivation and Positive Impact of Assessment	30	Assessment is a positive force for improving the social climate in a class.
	31	Classroom assessments motivate students to do their best.
Assessment as an Enjoyable Experience	32	Assessment is an engaging and enjoyable experience for students.

Comparison of Expected and Actual Factors.

This study explored the underlying dimensions of assessment-related beliefs and practices among educators. It compared theoretically expected factor groupings with those from our data analysis. Table 6 summarizes these comparisons, highlighting both alignments and discrepancies between the anticipated and actual factors.

Table 6: Expected Factors and Actual Factors

Expected Factor (Item)	Actual Factor (Item)	Comments
Assessment Purpose (1-6)	Assessment Purpose (1-8)	The inclusion of items 7 and 8 suggests that the actual data supports a broader conceptualization of assessment purpose than initially theorized. These additional items may also relate to how assessment is perceived in terms of its goals or intentions, leading to their inclusion in this factor.
Assessment Process (7-12)	Assessment Process (9-14)	This shift indicates that items 7 and 8, expected to belong here, aligned better with the Assessment Purpose factor. Items 13 and 14, which were not initially expected in this group, seem to align more with the practical aspects of the assessment process in the data.
Assessment Fairness (15-18)	Assessment Fairness (15-17)	This divergence suggests that fairness is perceived differently in practice, with items 17 and 18 potentially relating more closely to other factors such as Assessment Theory.
Assessment Theory (19-24)	Assessment Theory (18-21)	The split indicates that the theory-related items may involve distinct theoretical constructs. For instance, items 18 through 21 might relate to foundational theories, while 22 through 24 could pertain to specific applications or implications of those theories.

Table 6 (cont.)

	Assessment Theory (22-24)	This separation from the earlier grouping within Assessment Theory suggests that these items may represent a distinct aspect of theoretical understanding, potentially related to specific applications of assessment theory.
Assessment Beliefs (25-32)		The actual factors reveal a more nuanced understanding of assessment beliefs, suggesting that participants differentiate between negative, positive, and motivational aspects of assessment. The final item (32) stands out as its own factor, likely due to its distinct focus on the enjoyment aspect of assessments.
	Negative perceptions of assessment (25-27)	The actual factors reveal a more nuanced understanding of assessment beliefs, suggesting that participants differentiate between negative, positive, and motivational aspects of assessment.
	Time and overuse of assessments (28-29)	The grouping of these items indicates concerns specifically related to the time-consuming nature and potential overuse of assessments, a focus not as distinct in the original expected factor.
	Motivation and Positive Impact of Assessment (30-31)	This factor highlights positive beliefs regarding assessment's role in motivation and social climate, expanding the expected beliefs factor into a more positive dimension.
	Assessment as an enjoyable experience (32)	The final item stands out as its own factor, likely due to its distinct focus on the enjoyment aspect of assessments, which was not originally anticipated as a separate construct.

The factor analysis revealed insightful deviations from the expected structure of the assessment items. Initially, it was assumed that the items would align with predefined factors based on theoretical assumptions drawn from existing literature. However, the data suggested a different configuration, indicating the need to reevaluate how these factors are conceptualized across cultures. Barnes and colleagues (2017) state that teachers' conceptions of assessment differ across contexts, reflecting teachers' internalization of their cultural priorities and practices.

One of the key findings was related to the Assessment Purpose factor. While items 1 through 6 were expected to align with this factor, the analysis showed that items 7 and 8 also clustered under this category. This suggests that participants perceive the purpose of assessment in a broader context than initially theorized. These additional items may reflect a more comprehensive understanding of assessment goals, highlighting the importance of considering practical perceptions when defining the purpose of assessments.

The Assessment Process factor also revealed an unexpected configuration. Initially, items 7 through 12 were anticipated to belong to this factor. However, the analysis showed that items 7 and 8 aligned more with the Assessment Purpose than the Assessment Process. In contrast, items 13 and 14, not expected to be part of this factor, emerged as critical components of the Assessment Process. This finding indicates that participants may view the assessment process more practically, emphasizing aspects not initially considered in the theoretical framework.

Another interesting result was observed in the Assessment Fairness factor. The expectation was that items 15 through 18 would align with this factor. However, only items 15 and 16 matched this expectation, while items 17 and 18 were found to relate more closely to other

factors, such as Assessment Theory. This suggests that fairness is perceived more specifically by participants, reflecting distinct aspects of fairness that were not captured by the initial grouping.

The analysis of the Assessment Theory factor revealed a more complex structure than initially anticipated. The expected grouping included items 19 through 24, but the factor analysis split these into two groups. Items 18 through 21 appeared to relate to foundational theoretical constructs, while items 22 through 24 pertain to specific applications or implications of these theories. This split highlights the nuanced understanding of assessment theory among participants and suggests that different theoretical aspects are perceived as separate entities.

Lastly, the factor analysis of Assessment Beliefs uncovered a more detailed structure within this category. Although it was expected to encompass items 25 through 32, the analysis revealed distinct sub-factors. These included negative perceptions of assessment, concerns about time and overuse, and positive beliefs regarding assessment's motivational impact and social benefits. Interestingly, item 32 emerged as its factor, likely due to its focus on the enjoyment aspect of assessment, which was not anticipated in the initial theoretical model.

Overall, the factor analysis results suggest that the participants' perceptions of assessment are more complex and varied than the original theoretical model proposed. The emergence of new factors and the realignment of items within existing factors indicate a need for further exploration and refinement of these constructs in both theory and practice. This revised understanding could lead to more targeted and effective approaches to assessment in educational settings.

Composite scores for each dimension were calculated by averaging the items that load onto each factor, representing the participant's level of agreement or belief about each factor.

Reliability Analysis.

Cronbach's Alpha was calculated to assess the internal consistency of the items within each identified factor. Cronbach's Alpha is a measure of reliability that indicates how well the items within a factor are correlated and whether they consistently measure the same underlying construct. Table 7 shows the reliability analysis.

Table 7: Reliability Analysis

Expected Factor (Item)		Actual Factor (Item)		Note
Factor	Cronbach's Alpha	Factor	Cronbach's Alpha	
Assessment Purpose (1-6)	0.563	Assessment Purpose (1-8)	.651	The Cronbach's Alpha for the actual factor (1-8) is higher than expected (1-6), indicating improved internal consistency when additional items are included. This suggests that items 7 and 8, which were initially not included in the expected factor, contribute positively to the coherence of the assessment purpose factor.
Assessment Process (7-12)	.675	Assessment Process (9-14)	.688	The Cronbach's Alpha for the actual factor (9-14) is slightly higher than the expected factor (7-12), reflecting better internal consistency. This may be due to the inclusion of items 9-14, which align well with the practical aspects of the assessment process, leading to a more reliable factor.

Table 7 (cont.)

Assessment Fairness (13-18)	.589	Assessment Fairness (15-17)	.472	The actual Cronbach's Alpha for Assessment Fairness (15-17) is lower than the expected factor (13-18). This decrease suggests that the items 15-17 may not fully capture the intended concept of fairness as effectively as originally anticipated. Items 13 and 14, which were not included in the actual factor, might have contributed to the higher reliability observed in the expected factor.
Assessment Theory (19-24)	.699	Assessment Theory (18-21)	.756	The actual Cronbach's Alpha for Assessment Theory (18-21) is higher than the expected factor (19-24), indicating better internal consistency with this subset of items. This suggests that items 18-21 are particularly effective at capturing the theoretical aspects of assessment. However, items 22-24, which were not included in this factor, have a lower Cronbach's Alpha (0.461) in a separate factor, indicating a weaker reliability in capturing theory-related constructs.
		Assessment Theory (22-24)	.461	The lower Cronbach's Alpha for this group of items suggests that items 22-24 may be less reliable in measuring the theoretical constructs compared to items 18-21. This might indicate that these items pertain to different aspects of assessment theory that require separate consideration.
Assessment Beliefs (25-32)	.647			The Cronbach's Alpha for the expected factor (25-32) is 0.647, but specific sub-factors show varying reliability. This suggests that the overall reliability of assessment beliefs is acceptable, but further analysis of individual sub-factors is necessary.
		Negative perceptions of assessment (25-27)	.604	The Cronbach's Alpha for this sub-factor indicates moderate reliability in measuring negative perceptions of assessment. This suggests some consistency in responses regarding negative views but also highlights areas for potential improvement.
		Time and overuse of assessments (28-29)	.705	The higher Cronbach's Alpha for this sub-factor indicates good reliability in capturing views on time and overuse of assessments. This suggests a strong internal consistency among items related to this issue.
		Motivation and Positive Impact of Assessment (30-31)	.653	The reliability of this sub-factor is satisfactory, indicating that the items effectively measure motivation and the positive impact of assessment. However, there is room for improvement compared to the higher reliability observed in other sub-factors.
		Assessment as an enjoyable experience (32)		The Cronbach's Alpha for this single item is not provided, suggesting that its reliability should be evaluated in the context of overall assessment beliefs or potentially through further validation methods

Descriptive Analysis.

Table 8 presents the descriptive statistics for the composite scores of the identified factors, summarizing the mean, standard deviation (S.D), standard error (S.E), skewness, and kurtosis for each factor.

Table 8: Descriptive Statistics

Dimension	N	Mean (S.D)	Skewness (S.E)	Kurtosis (S.E)
Assessment Process	195	4.39 (.69)	-.344 (.174)	-.013 (.346)
Assessment Fairness	290	4.80 (.82)	-.591 (.143)	-.136 (.285)
Assessment Theory	300	4.59 (.79)	-.898 (.141)	1.967 (.281)
Assessment Purpose	289	4.68 (.56)	-.364 (.143)	1.120 (.286)
Assessment Theory 2	276	4.11 (.79)	-.427 (.147)	.605 (.292)
Negative Perspective about Assessment	296	3.10 (.99)	.083 (.142)	-.157 (.282)
Time and Overuse of Assessment	291	3.53 (1.22)	-.086 (.143)	-.634 (.285)
Motivation and Positive Impact	291	4.36 (.90)	-.483 (.143)	.196 (.285)

The descriptive statistics for the factors reveal varying levels of agreement among participants. Skewness values range from -0.898 to 0.083, and kurtosis values range from -0.157 to 1.967, indicating that the distributions are mostly within acceptable limits for normality (skewness: ±2, kurtosis: ±2). The mean scores suggest that participants generally leaned toward agreement across most factors, with the highest mean observed for "Assessment Fairness" and the lowest for “Negative Perceptions of Assessment.”

The ANOVA Analysis.

The ANOVA analysis in Table 9 evaluates the variability in several factors across different groups, particularly nine colleges. This statistical method examines how group means differ on various assessment factors, highlighting the differences in perceptions or experiences related to these factors. The table provides a comprehensive overview of the variability both between groups and within each group, offering insights into how these factors are perceived differently across the participating colleges. The following sections interpret the results for each factor, focusing on the significance of the differences observed.

Table 9: The ANOVA Analysis

Dimension	Group	Sum of Squares	df	Mean Square	F	Sig.
Assessment Process (AP)	Between Groups	10.734	8	1.342	3.053	.003
	Within Groups	81.753	186	.440		
	Total	92.487	194			
Assessment Fairness (AF)	Between Groups	28.353	8	3.544	5.994	<.001
	Within Groups	166.141	281	.591		
	Total	194.494	289			
Assessment Theory 1(AT1)	Between Groups	5.671	8	.709	1.141	.336
	Within Groups	180.856	291	.621		
	Total	186.527	299			
Assessment Purpose (AP)	Between Groups	1.577	8	.197	.615	.765
	Within Groups	89.725	280	.320		
	Total	91.302	288			
Assessment Theory 2 (AT2)	Between Groups	6.317	8	.790	1.290	.249
	Within Groups	163.423	267	.612		
	Total	169.740	275			
Negative Perspective about Assessment (NPA)	Between Groups	20.022	8	2.503	2.654	.008
	Within Groups	270.698	287	.943		
	Total	290.721	295			

Table 9 (cont.)

Time and Overuse of Assessment (TOA)	Between Groups	13.006	8	1.626	1.103	.361
	Within Groups	415.716	282	1.474		
	Total	428.722	290			
Motivation and Positive Impact (MPI)	Between Groups	11.829	8	1.479	1.888	.062
	Within Groups	220.894	282	.783		
	Total	232.723	290			

The ANOVA results revealed significant differences across groups for several factors related to assessment. Specifically, a statistically significant variation in perceptions of the Assessment Process, $F(8,186)=3.053, p=.003$, indicating that different groups perceive the assessment process differently. Additionally, perceptions of Assessment Fairness varied significantly across groups $F(8,281)=5.994, p<.001$, suggesting fairness is not uniformly perceived. Finally, significant differences were found in Negative Perceptions of Assessment (NPA), $F(8,287)=2.654, p=.008$, highlighting that some groups hold more negative views on assessment than others. These findings point to notable differences in how assessment is experienced and perceived across the studied groups.

In the robust tests of equality of means, the Welch test was used to account for potential violations of the assumption of homogeneity of variances. The results indicated significant differences in the means between groups for Assessment Process, $F(8,55.765)=3.444$, $p=.003$, and Assessment Fairness, $F(8,99.921)=5.675$, $p<.001$, suggesting that these factors are perceived differently across the groups studied. Additionally, significant differences were observed for Negative Perceptions of Assessment (NPA), $F(8,102.783)=2.691$, $p=.010$, indicating variability in negative views toward assessment among the groups. However, no significant differences were found for Assessment Theory, Assessment Purpose (APur), Assessment Theory 2 (AT2), Time and Overuse of Assessments (TOA), or Motivation and Positive Impact of Assessment (MPA), implying consistent perceptions across groups for these factors. These findings underscore the importance of understanding group-specific perceptions, particularly regarding the assessment process, fairness, and negative assessment perceptions.

The post-hoc Bonferroni tests, conducted following significant ANOVA results, revealed specific group differences in assessment perceptions (see Table 10). The GCIT group scored significantly higher in Assessment Fairness (AF) than the CLCS and SCE groups. Similarly, JNCE scored higher in Assessment Fairness (AF) than CLCS and SCE and higher in Assessment Purpose (AP) than GCB. The PCE group exhibited significantly higher scores in Assessment Purpose (AP) compared to both GCB and GCIT and in Assessment Fairness (AF) compared to CLCS and SCE. These findings indicate considerable variations in how different groups perceive the fairness and purpose of assessments.

Table 10: The Bonferroni Tests

Group Mean Diff (Sig.)	JNCE	PCE	SC
CLCS	.676 (.010) (AF)		
GCB	.732 (.001) (AF)		.60 (.041) (AP)
GCIT		-.972 (.037) (AP)	-.991 (.004) (AF)
		-.868 (.015) (AF)	
SC Mean	-.970 (.001) (AF)	-.742 (.006) (AP)	

* The mean difference is significant at the 0.05 level.

Conclusion

This study highlights the process of contextualizing an existing tool, the Approaches to Classroom Assessment Inventory Version 3 (ACAI-V3), to better understand faculty assessment beliefs at the Royal University of Bhutan. The contextualization process focused on two primary themes: adapting the ACAI-V3 (CART, 2019) dimensions to the local context and uncovering actionable insights that can guide the design of professional development (PD) initiatives (Malicka et al., 2019; Sims et al., 2020). The findings of this study provide a critical perspective on how faculty members perceive assessment, as well as the challenges they face when implementing assessment practices in their specific institutional and cultural context.

The ACAI-V3, originally structured around global theoretical constructs such as consistency, fairness, and balance, was designed to assess faculty beliefs on assessment. The tool's Part C focuses on four primary dimensions: assessment purpose, assessment process, assessment fairness, and assessment theory. However, the study revealed that these global constructs did not fully align with the realities of faculty experiences at the Royal University of Bhutan. To address this, the study employed a process of contextualization to adapt the tool and ensure it reflected local beliefs and practices (Behar-Hosenstein et al., 2014).

One of the key themes that emerged during this contextualization process was the adaptation of the dimensions of assessment purpose and fairness. The original ACAI-V3 categorizes fairness into standards, equity, and differentiation. However, faculty members at the Royal University of Bhutan indicated that their perceptions of fairness were shaped by local institutional policies, cultural norms, and specific challenges encountered during assessment implementation. In particular, fairness was often seen through the lens of aligning grading criteria with the diverse educational backgrounds of students, which presented additional challenges not captured by the original tool. Therefore, the contextualization of fairness required a deeper exploration of how these local factors influenced faculty beliefs, leading to a refined understanding of fairness within the Bhutanese educational context. Similarly, the dimension of assessment purpose, which traditionally includes assessment of learning, for learning, and as learning, was also adapted based on local practices. Faculty responses revealed a nuanced understanding of these purposes, where certain assessment functions were emphasized over others depending on the institutional context and teaching practices. This finding necessitated a revision of the assessment purpose dimension to capture better how faculty in Bhutan perceived the role of assessment in teaching and learning. Another significant area for adaptation was the dimension of assessment theory, particularly regarding consistency and contextualization. While the original ACAI-V3 strongly emphasizes consistency within assessment frameworks, the study identified that faculty at the Royal University of Bhutan faced significant challenges when trying to apply these theoretical frameworks in practice. This was primarily due to lacking resources, support, and clear guidance in implementing theoretical concepts in real-world assessments. Therefore, the tool was adjusted to reflect better the limitations faculty faced in applying theoretical constructs, particularly the need for more practical support in aligning assessment theory with local conditions.

The findings from this contextualized application of the ACAI-V3 tool emphasize the importance of tailored PD initiatives that address faculty members' specific needs and perceptions. The study revealed significant variation in faculty beliefs about assessment across different colleges within the Royal University of Bhutan. For example, some colleges

prioritized fairness and transparency, while others were more concerned with understanding the purpose of assessment. These findings underscore the inadequacy of a one-size-fits-all approach to PD (Behar-Hosenstein et al., 2014; Darling-Hammond et al., 2017; Malicka et al., 2017). Instead, PD programs should be customized to meet the distinct needs of different faculty groups, focusing on areas such as student-centered assessment strategies, innovative assessment tools, and enhancing faculty capacity to use assessment data for instructional decisions. The study also uncovered specific gaps in faculty practices related to negative perceptions of assessment, workload, and motivation. For instance, negative perceptions of assessment were often linked to high workloads, unclear assessment criteria, and perceived unfairness in grading. These issues pointed to a need for PD programs that address these concerns by providing strategies for managing assessment workloads, clarifying grading criteria, and improving fairness in assessment practices. Faculty members desired greater transparency in assessment processes, further reinforcing the need for PD programs to focus on these aspects.

Building on the insights gained from this study, several key areas emerge for future research and development to further refine assessment practices at the Royal University of Bhutan and similar institutions. One critical direction is developing context-specific professional development (PD) programs. This study identified significant gaps in faculty understanding of assessment fairness, the alignment of grading criteria with diverse student backgrounds, and the practical application of assessment theories. Tailored PD initiatives should be created to address these needs, focusing on enhancing faculty competence in these areas. Furthermore, continuous evaluation of PD programs will be necessary to ensure they remain relevant and practical, adapting to faculty members' changing needs and contexts over time (Desimone, 2009; Kohan et al., 2023). A second important avenue for future research involves cross-institutional and cross-cultural comparisons of assessment practices. The contextualization process used in this study offers an opportunity to explore how assessment tools, such as the ACAI-V3, can be adapted across different educational settings and cultural environments. By conducting comparative studies, researchers can better understand how universal constructs such as assessment fairness, purpose, and process might vary in different institutional contexts and the specific challenges faculty members face in each setting. This research could lead to broader insights into the flexibility and applicability of assessment frameworks.

Future studies could expand by incorporating student perspectives to provide a more holistic understanding of assessment practices. Comparing faculty and student perceptions of fairness, effectiveness, and alignment with learning objectives could uncover misalignments and improve assessment methods, ensuring they are fair, practical, and meaningful for all stakeholders. Additionally, exploring alternative assessment methods, such as formative, peer, and self-assessments, could address fairness concerns, reduce faculty workload, and increase student engagement. Research could also investigate how institutional policies shape faculty beliefs and practices, particularly regarding consistency and fairness, which could inform policies for more equitable assessment. Longitudinal studies could assess the long-term impact of contextualized assessment practices and PD interventions, identifying the most effective components. As technology becomes increasingly integrated into education, future research could explore how digital tools can streamline assessment, provide real-time feedback, and address fairness issues, modernizing assessment practices. Lastly, fostering collaboration among faculty in designing and reviewing assessment policies could reduce resistance to change, ensuring that assessment guidelines are practical and grounded in teaching realities and promoting ownership and accountability.

Acknowledgments

This study is part of a more extensive needs assessment conducted by the EUFSTIAT project team, which the Erasmus Plus Programme funds.

References

Barnes, N. Fives, H. & Dacey, C. M. (2017). U.S. teachers' conceptions of the purposes of assessment. *Teaching and Teacher Education, 65*, 107-116.

Bastos JL, Duquia RP, González-Chica DA, Mesa JM, Bonamigo RR. (2014). Fieldwork I: Selecting the instrument for data collection. *An Bras Dermatol*, *89*(6), 918-23. doi:10.1590/abd1806-4841.20143884.

Behar-Horenstein, L. S.; Garvan, C. W.; Catalanotto, F.A.; Hudson-Vassell, C.N. (2014). The role of needs assessment for faculty development initiatives. *Journal of Faculty Development*, *28*(2), 75-86.

Bergmark, U. (2020). Teachers' professional learning when building a research-based education: context-specific, collaborative and teacher-driven professional development. *Professional Development in Education, 49*(2), 210–224. https://doi.org/10.1080/19415257.2020.1827011

Bin, O. M., Mulyata, J., Meshari, A., & Debrah, Y. (2022). The challenges confronting the training needs assessment in Saudi Arabian higher education. *International Journal of Engineering Business Management, 14.* doi:10.1177/18479790211049706

CART. (2019). ACAI: Approaches to classroom assessment inventory: A tool for professional learning and assessment research. https://www.cdeluca.com/teacher-assessment-literacy

Creswell, J. W. (2019). Mixed methods and survey research in family medicine and community health. *Family Medicine and Community Health, 7*(2), e000086. https://doi.org/10.1136/fmch-2018-000086

Creswell, J. W. (2021). *A concise introduction to mixed methods research*. Sage Publications.

Darling-Hammond, L., Hyler, M. E., Gardner, M. (2017). Effective teacher professional development. Learning Policy Institute. https://doi.org/10.54300/122.311

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward Better Conceptualizations and Measures. *Educational Researcher, 38*(3), 181-199. https://doi.org/10.3102/0013189X08331140

James Jacob, W., Xiong, W. & Ye, H. (2015). Professional development programmes at world-class universities. *Palgrave Commun, 1,* 15002. https://doi.org/10.1057/palcomms.2015.2

Kohan, M., Changiz, T., & Yamani, N. (2023). A systematic review of faculty development programs based on the Harden teacher's role framework model. *BMC Med Educ 23,* 910. https://doi.org/10.1186/s12909-023-04863-4

Malicka, A., Gilabert Guerrero, R., & Norris, J. M. (2019). From needs analysis to task design: Insights from an English for specific purposes context. *Language Teaching Research, 23*(1), 78-106. https://doi.org/10.1177/1362168817714278

Sims, S., & Fletcher-Wood, H. (2020). Identifying the characteristics of effective teacher professional development: a critical review. *School Effectiveness and School Improvement, 32*(1), 47–63. https://doi.org/10.1080/09243453.2020.1772841

Steinert, Y., Mann, K., Anderson, B., Barnett, B. M., Centeno, A., Naismith, L., … Dolmans, D. (2016). A systematic review of faculty development initiatives designed to enhance teaching effectiveness: A 10-year update: BEME Guide No. 40. *Medical Teacher, 38*(8), 769–786. https://doi.org/10.1080/0142159X.2016.1181851

van Dijk, E. E., Geertsema, J., van der Schaaf, M. F., van Tartwijk, J., & Kluijtmans, M. (2023). Connecting academics' disciplinary knowledge to their professional development as university teachers: A conceptual analysis of teacher expertise and teacher knowledge. *Higher Education, 86*(4), 969-984.

Contact email: gembotshering.pce@rub.edu.bt