Curriculum Reform in Data Science Education: Enhancing Learning Outcomes With Scaffolding Learning Through Data Storytelling (SLDS)

Yiyun Fan, National University of Singapore, Singapore
Kah Loon Ng, National University of Singapore, Singapore
Wan Mei Amanda Soon, National University of Singapore, Singapore

The Asian Conference on Education 2024
Official Conference Proceedings

Abstract
Data storytelling (DS) employs narrative and visualization techniques to communicate insights from data, offering potential benefits for educational settings. This study introduces the framework of "Scaffolding Learning through Data Storytelling (SLDS)" as an explanatory approach to enhance students' learning outcomes in an undergraduate general education course on data literacy. Building on the key DS principles from Ryan (2016) and Knaflic (2015), we created a series of data stories to address students' diverse challenges in learning data science, taking into account their varied academic backgrounds, including STEM and non-STEM disciplines and differences in academic years. Incorporating Hadwin and Winnie's (2001) concept of "tacit scaffolds," SLDS integrates these stories into the curriculum stage by stage, aiming to enhance student engagement and understanding by encouraging them to read and think without explicitly directing or instructing specific studying activities. The effectiveness of SLDS was assessed through students' self-reported metrics of learning attention, relevance, confidence, and satisfaction, as well as multiple-choice questions measuring content comprehension. We anticipate that SLDS will improve learning outcomes more effectively than traditional methods, providing insights into easy-to-approach data narrative structuring and visualization design and its educational benefits for students from all backgrounds. This study aims to offer evidence on the application of DS in teaching and learning, laying a foundation for incorporating DS techniques into curricula and informing future educational practices for various educational levels and disciplines.

Keywords: Data Storytelling (Ds), Data Science Education, Curriculum Reform, Scaffolding in Learning, Learning Perception, Learning Comprehension

Introduction

Data science has emerged as a prominent discussion topic in education, with its widespread applications evident in everyday life. It touches nearly every discipline and professional field that involves working with large datasets. At its core, "data literacy," by its simplest definition, refers to a learner's ability of analyzing and communicating data. As society becomes more data-driven, the importance of developing data literacy skills has gained significant attention. In higher education, for instance, this has led to the introduction of numerous courses and degree programs focused on data science and statistics. By 2016, over 200 institutions had launched data science-related programs (De Veaux et al., 2017). Despite this progress and attention, there remains a significant shortage of professionals trained in data literacy to meet the growing demand in the workforce (Deja et al., 2021). Consequently, data science education still has a long way to go.

The swift advancements in educational technologies have equipped researchers with a myriad of tools to collect and analyze extensive datasets from learners, aiming to enhance and optimize the learning process through a practice known as learning analytics (Nunn et al., 2016; Yap et al., 2022). Many scholarly works have highlighted the importance of tracking and analyzing factors that can affect students' learning to improve overall teaching and curriculum design (LAK, 2011). However, in today's data science education, scholars continue to encounter various challenges in enhancing students' learning, especially at the curriculum level (Cassel & Topi, 2015). Efforts have been made to design and implement formal curricula that can make a meaningful impact on students across different age groups and within higher education settings (e.g., Dierker et al., 2017; Li et al., 2023). As a result, exploring ways to integrate different types of learning analytics into refining and advancing data science curricula remains essential, addressing both educational goals and societal demands.

In face of a large-context course with diverse student populations, effectively updating data science curricula, enhancing data literacy, and leveraging learning analytics to improve teaching and learning remain as significant challenges. One key challenge is addressing the varying needs of students from STEM and non-STEM backgrounds. While students from STEM fields typically receive more specialized training in statistics and mathematics, it is increasingly important to bridge the gap between these diverse groups. For instance, the concept of "creative data literacy" was introduced to engage students from non-STEM background in data science education (D'Ignazio, 2017). In Kim et al.'s (2024) study, they came up with a scale development process (Devellis, 2017; Boateng et al., 2018) to identify essential data literacy competencies. Additionally, Lim et al.'s (2021) study highlighted that the increasing number of students and the diversity of student population impact how learning analytics are applied in self-regulated learning (SRL).

Therefore, innovative approaches in data science education are crucial to facilitate active learning, especially in diverse educational and society contexts where students come from varying backgrounds but share a common goal of enhancing their data literacy skills. By incorporating new teaching and learning strategies into the curriculum, we can better address the needs of both STEM and non-STEM students, bridging knowledge gaps and supporting individualized learning. This approach not only improves learning outcomes and academic performance but also empowers students to regulate their own learning, preparing them for success in an increasingly data-driven world.

Data storytelling (DS), by its literal meaning, is a method of communicating insights from data using storytelling elements. DS emerges as a promising approach to address some of these challenges. DS, characterized by its ability to compress information and convey key elements through narratives and data visualizations (Ryan, 2016), holds the potential for enhancing learning experiences. It has been reported by recent scholars that DS elements, albeit with limited pedagogical constructs, have a promising future in educational settings (e.g., Chen et al., 2019; Echeverria et al., 2018; Martinez-Maldonado et al., 2020). Building upon this foundation, our study proposes the framework of "Scaffolding Learning through Data Storytelling (SLDS)" as an explanatory approach to enhance students' learning outcomes in an undergraduate general education course. By integrating the idea of "scaffolding" into the learning process (Wood et al., 1976), we aim to create more engaging, real-world, and accessible SRL content that resonates with students from diverse academic backgrounds. This method not only helps bridge gaps between different student populations but also provides the necessary self-learning support to foster deeper understanding and retention of data science concepts.

At the end of the study, we aim to leverage SLDS to create impact on students' perception and comprehension of introductory data science knowledge and assist with their data literacy development. By integrating SLDS into a course's curriculum, we seek to provide students with a structured curriculum framework for engaging with data and extracting meaningful insights. At the end of this study, we hope to answer the following two Research Questions (RQs) from both theoretical and practical aspects:

(1) How can generic data storytelling (DS) elements be incorporated into a data science course's curriculum?
(2) Does this incorporated design create impact on students' learning outcomes as they engage throughout the course?

Data Storytelling: What is it?

The power of stories stems from their presence in every aspect of our daily lives. From a young age, we engage in storytelling, weaving narratives that begin, unfold, conclude, and are retold, drawing upon our senses and personal experiences. While stories are inherently captivating, their impact is amplified when combined with data, offering audiences the ability to "refer, remember, and learn from them and how they affect our actions" (Tversky, 2024, p. 20). The storytelling of data, therefore, revolves around the origin of the data, the intended audiences, and the methods of its delivery, shaping diverse interpretations and influencing people through the narratives constructed.

The technique of DS serves as an information compression technique designed to convey important insights to audience (Ryan, 2016). Rooted in classic Information Visualization (InfoVis) principles and narrative storytelling elements (Tufte & Schmieg, 1985), DS is inherently explanatory (Martinez-Maldonado et al., 2020), with the aim of explaining insights within data and the importance of them to the audience (Echeverria et al., 2018). Recent applications of DS span various domains, including presentation of data using visualizations (Knaflic, 2015), data journalism (e.g., Ojo & Heravi, 2018), and teaching practices (e.g., Echeverria et al., 2018), although with somewhat limited pedagogical frameworks.

Understanding the principles of DS is crucial for leveraging its potential in educational contexts. As identified by Ryan (2016) and Knaflic (2015), DS is goal oriented, drives an audience's focus of attention, relies on choosing appropriate visuals, and adheres to core InfoVis design principles.

"Scaffolding" in Learning

The concept of "scaffolding" was originally introduced by Wood et al. (1976) to describe facilitative tools and skills that foster learner autonomy. Recent scholars including Renninger and List (2012), have further expanded on this, defining it as "a sustained interactive process that involves the fading of assistance/gradual task modifications by an expert" (p. 2923). Scaffolding in learning operates reciprocally, aiming to provide support that enables learners to engage in tasks independently. This approach differs from one-time and directive feedback or resources, focusing instead on an ongoing process of feedback provision.

There are various categories of scaffolding in learning. Hannafin et al. (1999) state that they include: (1) contextual—hints within contexts, (2) metacognitive—support specific for a particular task, (3) procedural—recourses for aiding task completion, and (4) strategic—different techniques or models. Such categories offer a more distinguished lining of scaffolding for supporting learning in classes.

Further discussion by Hadwin and Winne (2001) suggests "tacit scaffolds"—embedded tools that help to "cue students to attend to aspects of their studying without explicitly directing or instructing those studying activities" (p. 322). The process includes task understanding, setting goals and planning, enacting study tactics and evaluating and adapting metacognition. This differs from explicit scaffolds that have been brought about, further supporting the development of SRL.

Scaffolding Learning Through Data Storytelling

Scaffolding Learning through Data Storytelling (SLDS) is a theoretical framework that integrates a data storytelling-driven SRL approach with the principles of scaffolding in learning. Aimed at fostering students' self-regulation in learning data science, SLDS is proposed to be implemented in introductory data science courses through two distinct stages.

In the first stage, we propose the creation of specific forms of data storytelling (e.g., stories, comics, videos) that embed key data science knowledge and concepts. Leveraging the concept of tactic scaffolds (Hadwin & Winne, 2001), this approach emphasizes non-teacher-directed SRL, using fictional characters or narrative subjects who collaboratively discuss and analyze datasets. Through implicit tasks—such as setting research goals, engaging in character-driven discussions, making decisions, and evaluating analysis outcomes—the data storytelling-driven materials should progressively deepen the complexity of data analysis. These materials should also incorporate visualization tools to enhance readability and engagement. Additionally, their development must consider the concept of metacognition to align with students' cognitive processes (e.g., logical thinking during analysis), motivation (e.g., the interests of the stakeholders, the importance of datasets), and emotional engagement (e.g., analysis conflicts, team discussion, debates, leadership), ensuring the content is both compelling and resonates with students' learning needs (Hannafin et al., 1999).

In the second stage, SLDS emphasizes on addressing the structured phases of study in a data science course. Recognizing the challenges students might face—such as unmet learning needs or feeling overlooked due to the large scale of the course or difficulties in academic backgrounds (e.g., STEM vs. non-STEM, year one vs. year four)—SLDS highlights the importance of SRL throughout the course. To accommodate these diverse needs, the integration of data storytelling-driven materials should align with the course's structure and progressively match its levels of difficulty. For instance, in a data science course with four chapters of lecture content, the data storytelling-driven SRL materials should be strategically distributed at intervals throughout the course. This approach allows students to self-regulate their learning and reinforce the knowledge acquired in earlier chapters (see Figure 1).

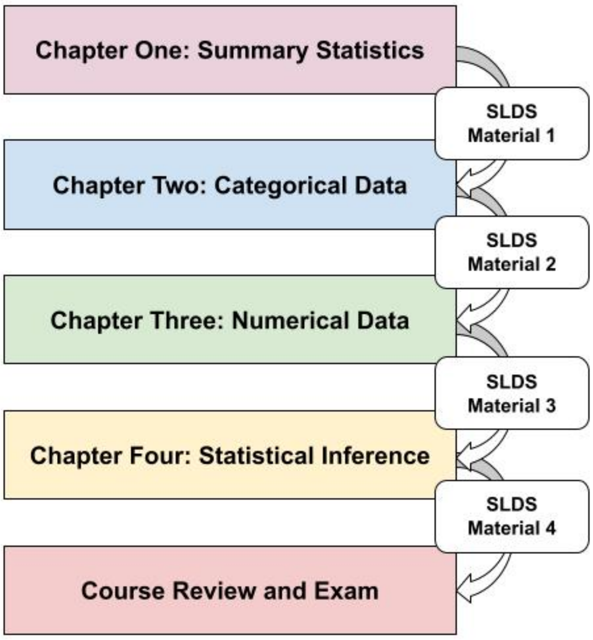


Figure 1: Example of SLDS Course Integration

Methodology

Participants

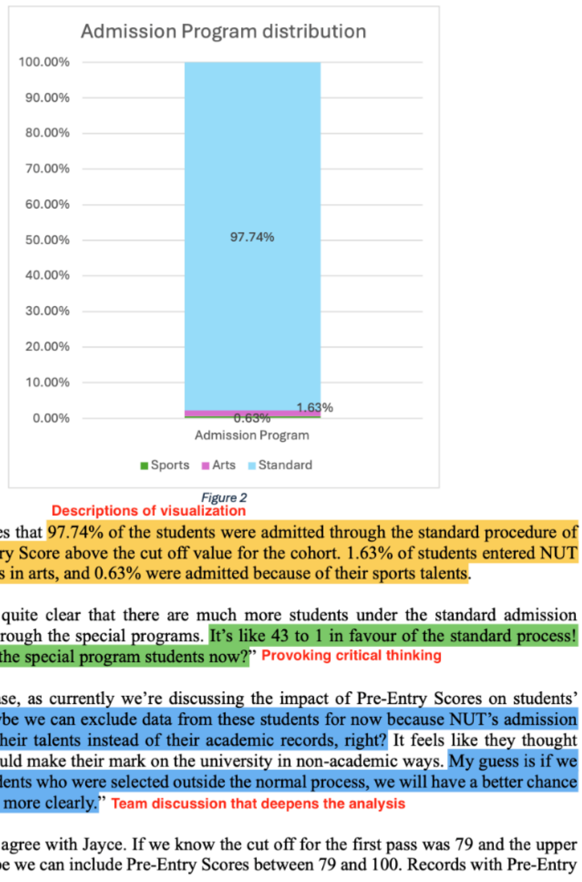
This study focused on students enrolled in an introductory data science course at a prestigious local university in Singapore. In the Academic Year 2023/2024 Semester 2, 2104 students were enrolled in the course, and 180 consented to provide access to their reported data. These students had no prior experience with data science-related courses at the university and were informed about the study's goal of enhancing their data science learning through a method called data storytelling.

The introductory data science course is a foundational entry-level program featuring a carefully curated curriculum designed by a teaching team of mathematicians and data scientists. Each semester, it attracts first- or second-year undergraduate students from diverse majors across the university. The course is designed to equip students with essential data literacy skills for analyzing data and making informed decisions in the face of uncertainty. It introduces basic principles and practice for collecting data and extracting useful insights, with examples drawn from various application domains (e.g., smoking and cancer correlations, housing market analysis). The data story materials used in this study were developed based on selected topics from the course, with the goal of supporting and enhancing students' learning experiences.

Data Story Materials and Evaluation Instrument

For the presentation of data storytelling-driven SRL materials, we opted to design four data stories that incorporated real-world scenarios and fictional characters within a data research team, simulating how data science projects or a set of research data unfold in real life. The design of the data stories followed the rules identified by Ryan (2016) and Knaflic (2015). There were four sets of stories in total, written and examined by a group of professionals, each with over five years of teaching experience in data science courses or expertise in pedagogical design.

These four data stories centered on a research team's efforts to analyze a dataset that included students' pre-entry scores at the university, their GPA trajectories over four years, and their subsequent employment outcomes and salaries. The narratives captured realistic actions and responses within a team-based data analysis setting, ensuring that the content was both engaging and authentic. Key concepts and knowledge points from the introductory data science course were seamlessly integrated into the stories, presented naturally through dialogues and visualizations to smoothen learning. Figure 2 showcases a visualization based on undergraduate admission data from a local university, accompanied by a narrative framework derived from the visualization. The narrative included a description of the visualization, questions posed by a fictional character to stimulate critical thinking, and a team discussion designed to deepen the analysis.



The chart indicates that 97.74% of the students were admitted through the standard procedure of earning a Pre-Entry Score above the cut off value for the cohort. 1.63% of students entered NUT due to their talents in arts, and 0.63% were admitted because of their sports talents.

Luke: "Now it's quite clear that there are much more students under the standard admission procedure than through the special programs. It's like 43 to 1 in favour of the standard process! Can we filter out the special program students now?"

Jayce: "In this case, as currently we're discussing the impact of Pre-Entry Scores on students' future GPAs, maybe we can exclude data from these students for now because NUT's admission office looked at their talents instead of their academic records, right? It feels like they thought these students would make their mark on the university in non-academic ways. My guess is if we exclude these students who were selected outside the normal process, we will have a better chance of seeing patterns more clearly."

Jennifer: "Well, I agree with Jayce. If we know the cut off for the first pass was 79 and the upper limit is 100, maybe we can include Pre-Entry Scores between 79 and 100. Records with Pre-Entry

Figure 2: Excerpt Extracted From One Data Story

Evaluation of the proposed approach included two main components—learning perception and learning comprehension. In the first component, participants were encouraged to evaluate their learning perception using a three-point Likert scale (yes, no, and uncertain) to indicate their perceptions. To interpret students' learning perceptions, we adopted Keller's (1987) and Keller and Kopp's (1987) ARCS framework, which classify learning evaluations into four dimensions: attention, relevance, confidence, and satisfaction. The index of attention reflects learners' interest in or motivation for learning. Relevance measures the extent to which a learner perceives the current content as connected to their prior knowledge or experiences. Learning confidence represents learners' positive expectations for achieving successful learning outcomes. Satisfaction evaluates whether the learning outcomes align with learners' expectations, indicating their ability to understand the material provided in this study's context and offer positive feedback. Previous studies have demonstrated that the ARCS model effectively tailors to students' needs and interests, stimulating attention, attracting interest, encouraging learning, and fostering satisfactory learning outcomes (e.g., Afjar et al., 2020; Karyani, 2017; Liu & Hou, 2021). To understand students' learning, the ARCS model was deemed most suitable for this study, allowing students to assess their own learning outcomes during the learning process. For instance, Figure 3 shows some examples of the learning perceptions questions from our evaluation survey.

Attention:

I want to learn more about the data story after reading it.

☐ Yes ☐ No ☐ Uncertain

Relevance:

I have referred back to course or external materials to refresh the statistical knowledge learned while reading the data story.

☐ Yes ☐ No ☐ Uncertain

Confidence:

I am confident that I am mastering the content in the data story.

☐ Yes ☐ No ☐ Uncertain

Satisfaction:

I understand the data story well.

☐ Yes ☐ No ☐ Uncertain

Figure 3: Examples of Learning Perception Questions

In the second component, we focused on a more objective evaluation of learning outcomes. Participants were tasked with reading the data stories and completing a series of multiple-choice questions (MCQs) and multiple-response questions (MRQs) to test their comprehension. These questions were carefully crafted to align with the knowledge presented in the stories, emphasizing key concepts and addressing areas that students might commonly misunderstand. By assessing participants' responses, these questions served as objective measures to gauge learners' comprehension and retention of the key concepts in the course introduced through the SLDS approach. Figure 4 is an example of a learning comprehension question.

The R value of Pre-Entry Score vs. Term 1 GPA is at 0.417 and the one of Pre-Entry Score vs. Final Term GPA is at 0.114. Based on what you have learned, what do these two figures mean to you? Select all that apply.

A. The strength of the linear relationship between the Pre-Entry Score and the Term 1 GPA is stronger than that between the Pre-Entry Score and the Final Term GPA.
B. R value at 0.114 indicates that Pre-Entry Score and Final Term GPA are almost unrelated to each other.
C. Correlation coefficient R shows the strength of the linear relationship between variables.
D. R value at 0.417 indicates that the Pre-Entry Score and Term 1 GPA are moderately associated with each other.

(Suggested answers: A, C, and D)

Figure 4: Example of a Learning Comprehension Question

Collectively, these evaluation methods provided comprehensive learning analytics insights into the effectiveness of SLDS in impacting on students' learning outcomes and promoting deeper understanding among learners. Future educators can use the student-reported analytics data to better tailor the content and context of SLDS, accommodating the diverse learning abilities of various student groups, especially in a large population.

Data Collection Procedures

The reflection surveys were distributed to students through the course's online learning system. For each set of data story materials, students were required to read the content and complete the corresponding reflection survey within a two-week period. Students were encouraged to read the data story first and then complete the survey to evaluate their understanding and learning. To encourage participation, students who completed each reflection survey received 0.5 points toward their final grade. After all responses for the four reflection surveys were collected, we assessed participant eligibility by verifying two criteria within the survey: (1) whether they had provided full consent for their survey data to be included in the study, and (2) whether they had acknowledged reading the full data story before proceeding to the survey.

Analysis

Learning Perceptions

In terms of learning perception questions related to learning attention, relevance, confidence, and satisfaction, we gave two marks if a participant indicated *Yes*, one mark for *Uncertain*, and 0 mark for *No*. Table 1 presents the average marks for each survey in terms of the four attributes of the 180 participants' learning perceptions.

Survey	Attention	Relevance	Confidence	Satisfaction
1	3.11	1.03	6.57	1.77
2	2.44	0.87	6.39	1.77
3	2.83	1.04	6.54	1.76
4	2.68	0.96	6.30	1.69

Table 1: Average Scores for Each Attribute Across All Four Surveys

Based on participants' average scores, we calculated the proportions of the 180 participants who scored above and below average, for each attribute in learning perception, across all four surveys. Table 2 summarizes the results. Overall, the majority scored above average in learning attention questions in all the surveys, with a slight dip in survey 2. Learning relevance showed a more balanced distribution, with surveys 2 and 4 slightly towards below-average scores. For learning confidence, the majority consistently scored above average, with percentages fluctuating between 57.78% and 62.78%. Lastly, learning satisfaction consistently had the highest above-average scores among the four attributes, peaking at 79.44% in survey 1.

	Attention		**Relevance**		**Confidence**		**Satisfaction**	
Survey	Above average	Below average	Above average	Below average	Above average	Below average	Above average	Below average
1	61.00%	39.00%	51.67%	48.33%	62.78%	37.22%	79.44%	20.56%
2	41.11%	58.89%	43.33%	56.67%	57.78%	42.22%	78.33%	21.67%
3	64.44%	35.56%	52.22%	47.78%	60.56%	39.44%	77.78%	22.22%
4	59.44%	40.56%	47.78%	52.22%	60.00%	40.00%	73.78%	27.22%

Table 2: Proportions of Participants Who Scored Above and Below Average for Each Survey

To deepen our analysis, we also performed an ANOVA test on R to see if there are statistically significant differences in the learning perception outcomes across the four surveys. We obtained an F-value of 6.405. Note that the F-value represents the ratio of the variance of the scores between the surveys to the variance of the scores within the surveys. The corresponding *p*-value of 0.000287 showed significant differences in scores between at least two of the surveys. This suggests that students' self-reported answers may have been influenced by the specific story and survey they completed.

Learning Comprehension

For learning comprehension questions, we designed multiple-response questions (MRQ) and multiple-choice questions (MCQ) that tested participants' understanding of the knowledge presented in the stories. We awarded two marks to a participant only if all the correct options for an MRQ were chosen, one mark for a partially correct answer, meaning the participant had chosen at least one correct option, and 0 mark if none of the correct options was chosen. For MCQs with only one correct option per question, we offered one mark for a correct answer and 0 mark for a wrong answer.

Table 3 presents the average marks for the learning comprehension questions, along with the proportions of the 180 participants who scored above and below average across all surveys. Over half of the participants scored above average in surveys one and two. However, as the complexity of the knowledge presented in the stories increased in later surveys, participants appeared to struggle with comprehending all the concepts. This likely contributed to generally lower scores in surveys 3 and 4.

Survey	Learning comprehension	Above average	Below average
1	3.633/5	55.56%	44.44%
2	5.610/7	63.89%	36.11%
3	4.120/6	32.78%	67.22%
4	3.122/6	30.00%	70.00%

Table 3: Average Score and Proportions of Participants Who Scored Above and Below Average for Each Survey

We conducted an ANOVA test on R to determine whether there were statistically significant differences in participants' learning comprehension across the four surveys. The analysis yielded an F-value of 152.809 and a *p*-value of 1.357625e-71 (essentially 0), indicating significant differences in scores between at least two of the surveys. This suggests that students' comprehensions may have been influenced by the variations in the type or depth of content presented in the different stories.

Conclusion

In this study, we aimed to achieve two key objectives: (1) to explore how generic DS elements can be scaffolded to enhance students' learning experiences, and (2) to assess the potential impact of DS on students' self-reported learning perceptions and objective comprehension. We focused on identifying effective strategies for structuring and presenting SRL materials in hope of improving students' engagement in learning. The feedback in learning perceptions and comprehension data collected from students served as valuable learning analytics, offering guidance for refining and optimizing the SLDS content in the future.

The significant differences between surveys in both learning perception and learning comprehension components showcase that SLDS could possibly have a positive impact on students' SRL of introductory data science knowledge. However, we were also cautious about the data collected due to several limitations. Firstly, as the data was self-reported by participants, the subjective nature of learning perception introduces some uncertainty regarding the impact. Secondly, with only 180 students participating in this voluntary study, the sample might not be fully representative given the large course enrollment. Thirdly, data storytelling-driven SRL materials are not limited to the format of traditional data stories. The technique of data storytelling allows for various forms of creative applications. However, with vague guiding principles and no specific instructions provided, these materials can take on different formats, making their implementation in SRL flexible yet potentially inconsistent.

Nevertheless, this study's investigation furnished practical evidence on the potential effectiveness of SLDS in curriculum and pedagogical practices, thus offering a pathway for future researchers and math educators to integrate DS techniques into their everyday teaching practices. Moreover, SLDS can provide insights into students' performance. The data obtained from evaluations (i.e., learning perception and comprehension) can be instrumental in upgrading future content and adapting teaching and learning varieties across different educational contexts (e.g., K-12 and other higher education contexts). By adapting scaffolded data stories, or other innovative formats of SRL materials, to cognitive processes and motivations, we provided insights into the potential benefits of SLDS framework in fostering a deeper understanding of the data knowledge and analytical processes.

Acknowledgement

This work was supported by the Singapore Ministry of Education under Singapore Ministry of Education Tertiary Education Research Fund (MOE2022-TRF-016).

References

Afjar, A. M., Musri, & Syukri, M. (2020). Attention, relevance, confidence, satisfaction (ARCS) model on students' motivation and learning outcomes in learning physics. *Journal of Physics: Conference Series, 1460*(1), 1–7. https://doi.org/10.1088/1742-6596/1460/1/012119

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Cassel, B., & Topi, H. (2015). Strengthening data science education through collaboration. Workshop report, NSF Award #DUE1545135. National Science Foundation.

Chen, Q., Li, Z., Pong, T.-C., & Qu, H. (2019). Designing narrative slideshows for learning analytics. In *Proceedings of the IEEE Pacific Visualization Symposium, PacificVis'19* (pp. 237–246). https://doi.org/10.1109/PacificVis.2019.00036

Deja, M., Januszko-Szakiel, A., Korycińska, P., & Deja, P. (2021). The impact of basic data literacy skills on work-related empowerment: The alumni perspective. *College & Research Libraries, 82*(5), 708–725. https://doi.org/10.5860/crl.82.5.708

De Veaux, R. D., Agarwal, M., Averett, M., Baumer, B. S., Bray, A., Bressoud, T. C., Bryant, L., Cheng, L. Z., Francis, A., Gould, R., Kim, A. Y., Kretchmar, M., Lu, Q., Moskol, A., Nolan, D., Pelayo, R., Raleigh, S., Sethi, R. J., Sondjaja, M., Tiruviluamala, N., Uhlig, P. X., Washington, T. M., Wesley, C. L., White, D., & Ye, P. (2017). Curriculum guidelines for undergraduate programs in data science. *Annual Review of Statistics and Its Application, 4*(1), 15–30. https://doi.org/10.1146/annurev-statistics-060116-053930

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4^{th} ed.). Sage

Dierker, L., Ward, N., Alexander, J., & Donate, E. (2017). Engaging underrepresented high school students in data-driven storytelling: An examination of learning experiences and outcomes for a cohort of rising seniors enrolled in the Gaining Early Awareness and Readiness for Undergraduate Program (GEAR UP). *Journal of Education and Training Studies, 5*(4), 54–63. https://doi.org/10.11114/jets.v5i4.2187

D'Ignazio, C. (2017). Creative data literacy: Bridging the gap between data-haves and data-have-nots. *Information Design Journal, 23*(1), 6–18. https://doi.org/10.1075/idj.23.1.03dig

Echeverria, V., Martinez-Maldonado, R., Shum, S. B., Chiluiza, K., Granda, R., & Conati, C. (2018). Exploratory versus explanatory visual learning analytics: Driving teachers' attention through educational data storytelling. Journal of Learning Analytics, 5(3), 72–97. http://dx.doi.org/10.18608/jla.2018.53.6

Karyani, Y. (2017). Implementasi model pembelajaran ARCS (Attention, Relevance, Confidence, Satisfaction) untuk meningkatkan minat dan hasil belajar matematika siswa kelas IX-E MTs Negeri 2 Mataram pada materi peluang tahun pelajaran 2016/2017. *Jurnal Ilmiah Mandala Education, 3*, 189. http://dx.doi.org/10.58258/jime.v3i1.36

Keller, J. M. (1987). Development and use of ARCS model of instructional design. Journal *of Instructional Development, 10*(3), 2–10. https://doi.org/10.1007/BF02905780

Keller, J. M., & Kopp, T. W. (1987). An application of the ARCS Model of Motivational Design. In C. M. Reigeluth (Ed.), *Instructional theories in action: Lessons illustrating selected theories and models* (pp. 289–320). Lawrence Erlbaum Associates, Inc.

Kim, J., Hong, L., & Evans, S. (2024). Toward measuring data literacy for higher education: Developing and validating a data literacy self-efficacy scale. Journal of the Association for Information Science and Technology, 75, 916–931. https://doi.org/10.1002/asi.24934

LAK. (2011). Proceedings of the 1st International Conference on Learning Analytics and Knowledge.

Li, Y., Wang, Y., Lee, Y., Chen, H., Petri, A. N., & Cha, T. (2023). Teaching data science through storytelling: Improving undergraduate data literacy. *Thinking Skills and Creativity, 48*, 1–12. https://doi.org/10.1016/j.tsc.2023.101311

Lim, L.-A., Gentili, S., Parfo, A., Kovanovic, V., Whitelock-Wainwright, A., Gasevic, D., & Dawson, S. (2021). What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learning and Instruction, 72*, 101202. https://doi.org/10.1016/j.learninstruc.2019.04.003

Liu, Y.-M., & Hou, Y.-C. (2021). Effect of multi-disciplinary teaching on learning satisfaction, self-confidence level, and learning performance in nursing students. *Nurse Education in Practice, 55*, 103128. https://doi.org/10.1016/j.nepr.2021.103128

Martinez-Maldonado, R., Echeverria, V., Nieto, G. F., & Shum, S. B. (2020). From data to insights: A layered storytelling approach for multimodal learning analytics. Paper presented at the *CHI '20 Conference on Human Factors in Computing Systems*, April 25-30, 2020, Honolulu, HI, USA.

Nunn, S., Avella, J., Kanai, T., & Kebritchi, M. (2016). Learning Analytics Methods, Benefits, and Challenges in Higher Education: A Systematic Literature Review. *Online Learning*, 20, 13–29.

Ryan, L. (2016). *The visual imperative: Creating a visual culture of data discovery*. Elsevier Science.

Wood, D., Bruner, J., & Ross, G. (1976). The role tutoring in problem solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 17*, 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Yap, S. M., Khalid, F., Wong, A. C. Q., & Ismail, S. S. (2022). Learning analytics in education, advantages and issues: A systematic literature review. *Creative Education, 13*(9), 2913–2920. http://doi.org/10.4236/ce.2022.139183

Contact emails: yiyunfan@nus.edu.sg
kloon@nus.edu.sg