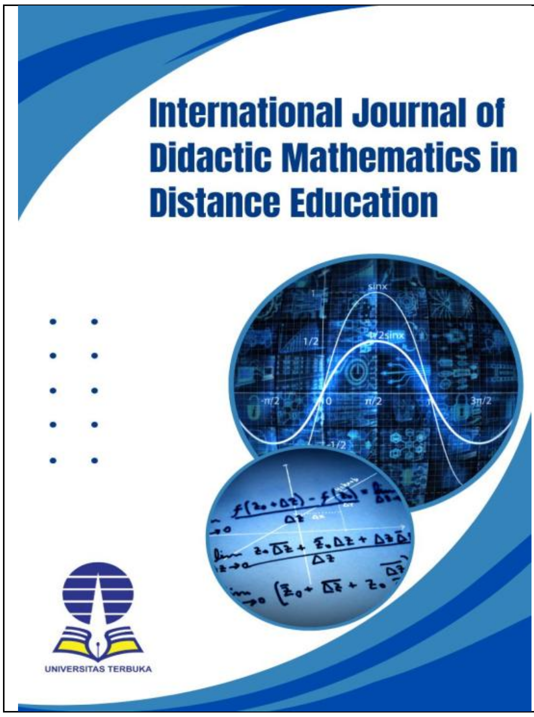
International Journal of Didactic Mathematics in Distance Education

Journal homepage: https://jurnal.ut.ac.id/index.php/ijdmde



Integrating technology and non-routine problems: the design, development, and validation of a computer-aided assessment for non-routine problem-solving

Marivic G. Molina[1,2], Erna V. Yabut[3]

[1]Graduate School, Centro Escolar Univerity, Manila, Philippines, molina1715373@ceu.edu.ph

[2]College of Science, Pamantasan ng Lungsod ng Maynila, Manila, Philippines, mgmolina@plm.edu.ph

[3]Graduate School, Centro Escolar Univerity, Manila, Philippines, evyabut@ceu.edu.ph

To cite this article:

Molina, M.G & Yabut, E.V. (2025). Integrating technology and non-routine problems: the design, development, and validation of a computer-aided assessment for non-routine problem-solving. *International Journal of Didactic Mathematics in Distance Education*, *2*(1), 81-104.

To link to this article:

https://jurnal.ut.ac.id/index.php/ijdmde

Published by:

Universitas Terbuka

Jl. Pd. Cabe Raya, Pd. Cabe Udik, Kec. Pamulang, Kota Tangerang Selatan, Banten 15437

International Journal of Didactic Mathematics in Distance Education
Volume 2, Issue 1, pp. 81-104, E-ISSN: 3047-9207
https://jurnal.ut.ac.id/index.php/ijdmde

Integrating technology and non-routine problems: the design, development, and validation of a computer-aided assessment for non-routine problem-solving

Marivic G. Molina[1,2*], **Erna V. Yabut**[3]
[1*]Graduate School, Centro Escolar University, Manila, Philippines, molina1715373@ceu.edu.ph
[2]College of Science, Pamantasan ng Lungsod ng Maynila, Manila, Philippines, mgmolina@plm.edu.ph
[3]Graduate School, Centro Escolar University, Manila, Philippines, evyabut@ceu.edu.ph

*Correspondence: mgmolina@plm.edu.ph

Abstract
The increasing demand for higher-order thinking skills in 21st-century education underscores the need for innovative resources, as traditional materials often focus on routine problems with limited opportunities for non-routine problem-solving. This study designed, developed, and validated a computer-aided assessment for non-routine problem-solving using the Design and Development Research method, conducted in three phases: needs analysis, design and development, and evaluation. Data collection included a literature review, systematic design and development, expert validation, and user feedback, analyzed through descriptive statistics and thematic analysis. Findings show that the computer-aided assessment supports non-routine problem-solving and addresses limitations of traditional assessments. The needs analysis identified gaps in current methodologies and the benefits of technology-enhanced assessments. The design and development phase produced an assessment promoting cognitive engagement, while the evaluation phase gathered expert validation and user feedback. Five experts in information technology and mathematics education provided positive evaluations of the assessment's functionality, accessibility, and cognitive support. Feedback from 15 students across three educational institutions indicated enhanced engagement, independent learning, and deeper mathematical understanding. Although minor technical issues were noted, they did not significantly impact the computer-aided assessment's quality. The validated computer-aided assessment shows potential as a digital tool for fostering non-routine problem-solving and interactive learning. It suggests that computer-aided assessments with non-routine problems may enhance higher-order thinking skills by promoting cognitive engagement and personalized learning. Future research could explore its long-term impact, adaptability across disciplines, and potential for adaptive learning and collaborative features to maximize educational effectiveness.

Article History
Received:
5 March 2025
Revised:
30 March 2025
Accepted:
09 April 2025
Published Online:
30 April 2025

Keywords:
Computer-aided assessment;
Design and Development Research;
Higher-order thinking skills;
Non-routine problems

1. Introduction

In the rapidly evolving landscape of education, integrating technology into teaching and learning has become increasingly essential for preparing students to meet the demands of the 21st century. The growing use of computers and the internet significantly transformed the way businesses and markets functioned, and the education sector was no exception (Bianchi et al., 2020). This transformation was further supported by Kartikasari et al (2022), who emphasized that technological advances in the 21st century influenced every field, including education. The COVID-19 pandemic underscored the importance of technology, with educational institutions worldwide relying on digital tools to maintain learning during in-person class disruptions.

This increasing reliance on digital tools highlighted the need to equip students with 21st-century skills like Higher-Order Thinking Skills (HOTS). These competencies encouraged students to engage in critical thinking, explore various approaches, and develop innovative solutions when faced with unfamiliar problems, questions, uncertainties, dilemmas, or new situations (Truong & Tran, 2021). This emphasis on HOTS was particularly evident in the context of non-routine problem-solving, which required students to apply their knowledge in unfamiliar and complex situations, making the use of HOTS essential (Makmuri et al., 2021; Shawan et al., 2021; Kablan & Uğur, 2019). Engaging in mathematical non-routine problems (NRPs) proved particularly valuable for developing these skills (Mogari & Lupahla, 2013). Andrade et al (2020) suggested that NRPs can be successfully solved by students of all ability levels in regular schools if they receive appropriate interventions and instruction in problem-solving techniques. However, NRPs were not frequently included in textbooks or instructional materials (Berisha, 2015; Fan & Zhu, 2000; Kablan & Uğur, 2019; Kolovou et al., 2009; Manopo & Lisarani, 2021; van Zanten & van den Heuvel-Panhuizen, 2018) due to the complexity involved in designing and assessing such problems, limiting students' opportunities to cultivate these essential competencies.

Traditional instructional resources often prioritized routine problems with clear, prescribed solutions, which were easier to grade and standardize. To address this issue, Pegg (2010) suggested that teachers present students with non-routine questions to promote HOTS development. Mikusa (Legarde, 2022) emphasized that enhancing students' HOTS required engagement with complex problems that could take hours, days, or even weeks to solve. Similarly, Barana et al (2019) highlighted the importance of regularly exposing students to a wide variety of problems over an extended period to effectively develop their problem-solving skills.

One of the key challenges in exposing students to NRPs was the need for timely and effective feedback, which was difficult to provide in traditional classroom settings. According to Van der Kleij et al (2015), feedback can assist students in recognizing and rectifying mistakes and misunderstandings, developing better and more efficient problem-solving approaches, and enhancing their self-regulation. When working with NRPs, students were encouraged to apply unconventional techniques, often leading to solutions that were not easily predictable (Kurniawan et al., 2022). Usta (2020) stressed the importance of students understanding how they learned most effectively and recognizing their mistakes to correct them—making immediate feedback crucial to the learning process. However, in classrooms with 30 to 40 students, providing immediate feedback to each student became difficult and required significant effort from teachers.

CAA addressed this challenge by streamlining distribution, minimizing assessment and grading time, automating test result reporting, and enhancing usability and accessibility for students (Dyrvold & Bergvall, 2023). Beyond logistical advantages, CAA also provides practical benefits, such as delivering immediate feedback on students' final answers (Rasila et al., 2015). In digital learning environments, feedback was delivered almost instantly, as it was automatically generated based on students' responses (Van der Kleij et al., 2015). Miller (2009, as cited in Van der Kleij et al., 2015) noted that recent studies showed students preferred immediate feedback over delayed responses. Furthermore, the automation of the scoring process promoted more independent learning, creating a more efficient and self-regulated learning environment (Barana et al., 2022).

Over the past decade, the availability of CAA has grown substantially (Davies et al., 2022; Kinnear et al., 2022), with the COVID-19 pandemic further intensifying the need for its adoption (Davies et al., 2022). Since 2020, higher education shifted to online learning (Su & Guo, 2021) and progressively adopted blended learning approaches (Lee et al., 2024) in response to the ongoing impact of the pandemic on global education systems (Saro et al., 2022). In the Philippines, this shift resulted in significant changes (Saro et al., 2022), including the adoption of blended learning approaches, the integration of digital tools for assessment and instruction, and a growing emphasis on independent and self-regulated learning. These changes reshaped teaching methods and learning outcomes, pushing both educators and students to develop new competencies suited to a technology-driven environment, although there was now a gradual return to fully face-to-face instruction.

While CAA had demonstrated promise in enriching learning experiences and providing immediate feedback, its application in NRP-solving had remained limited. This study aimed to bridge that gap by designing, developing, and validating a CAA specifically tailored for NRP-solving. Unlike previous efforts confined to traditional paper-based formats, this CAA functioned as a dynamic test bank for formative assessment, offering an interactive and adaptable platform. Building on this foundation, the primary objective of this study was to design, develop, and validate both the content and functionality of the CAA for solving NRPs through expert validation and user feedback. These processes ensured its quality, relevance, and usability while supporting deeper learning and fostering HOTS and problem-solving abilities. By incorporating usability principles, the CAA aimed to create an intuitive and engaging learning environment that minimized cognitive load and allowed students to focus on applying their HOTS and problem-solving strategies. To achieve this, the study addressed three key research questions:

a) What were the needs and requirements, based on the literature review, for developing a CAA for solving NRPs?
b) How was the CAA designed and developed to meet these needs?
c) How did experts and users evaluate the quality and usability of the CAA?

.

2. Method

2.1 Research Design

This study employed the Design and Development Research (DDR) method. According to Richey and Klein (2007), DDR involved a systematic investigation of the processes of design, development, and evaluation, aiming to build an empirical foundation for creating instructional and non-instructional products, tools, and improved or newly developed models that guided their creation. Similarly, Plomp (2007) argued that this approach was also used in designing and developing interventions such as programs, instructional and learning strategies, products, and systems. These interventions addressed complex educational challenges and enhanced the understanding of the characteristics and processes involved in their design and development.

The DDR method suited this study as it guided the design, development, and validation of a CAA for non-routine problem-solving. The study followed three phases: needs analysis, design and development, and evaluation, ensuring the systematic creation of the CAA. The needs analysis phase was the initial stage in DDR, conducted to determine the necessity for development (Hasbullah et al., 2022). The design and development phase is the most important phase in DDR (Jaya et al., 2021). According to Van Den Akker et al. (Padzil et al., 2021), this phase was essential and needed to be prioritized to ensure that the developed products — whether modules, models, or curricula — were relevant and carefully refined to effectively serve the intended target audience. Finally, Hasbullah et al. (2022) reported that during the evaluation phase, many studies on product and tool development employed both quantitative and qualitative methodologies.

2.2 Participants and Sampling

In DDR, Richey and Klein (2007) explained that product and tool studies typically involved designers, developers, clients, evaluators, and users. More specifically, when examining the impact of an innovative product or tool, it was crucial to include those directly affected by its use, such as clients, learners, instructors, and sometimes supervisors. This framework provided a clear understanding of the importance of involving diverse stakeholders in the evaluation process.

Applying this perspective, the study's participants consisted of five expert validators. These included three IT specialists—two of whom held Doctorates in IT and one with a master's degree—and two Mathematics and the Modern World (MMW) professors with master's degrees. The decision to choose only five experts was made to balance the need for diverse expertise with the practical considerations of time and resource constraints. The inclusion of these experts aligned with the recommendations of Ikart (2019), who noted that the number of expert reviewers typically ranged between three and five, though some studies had reported as many as 20 reviewers. The expert validators were selected through purposive sampling based on their expertise, their ability to evaluate the CAA, and their willingness to complete the evaluation on time. In purposive sampling, also called judgmental or expert sampling, participants were chosen by the researcher based on their suitability for providing the needed data (Ahmed, 2024).

In addition to the expert validators, fifteen students participated in exploring the CAA. These students were drawn from three institutions—five from a sectarian college, five from a local university, and five from a non-sectarian university. All were enrolled in MMW for SY 2023-2024 and were selected through convenience sampling, including only those who submitted the informed consent form, expressed an interest in exploring the CAA, and were willing to participate. Jhangiani et al. (2019) described convenience sampling as the most common alternative to random sampling. In this method, the sample consisted of individuals who were easily accessible and willing to take part in the study. This approach facilitated timely data collection while ensuring a diverse representation of students from different educational settings.

2.3 Data Collection

The data collection process involved gathering feedback from both experts and students to validate and refine the CAA. Two instruments were used: a rubric and a feedback form. Expert validation was conducted using the Rubric for e-Learning Tool Evaluation, adapted from Anstey and Watson (2018) work and shared under a Creative Commons license. This rubric provided a structured framework for assessing the CAA's content and functionality. The adapted Rubric for e-Learning Tool Evaluation was enhanced by incorporating numerical values to complement its verbal interpretations. To create a more refined Likert-scale approach, the following scoring ranges were established: 2.01-3.00 for "Works Well," 1.01-2.00 for "Minor Concerns," and 0.00-1.00 for "Serious Concerns." These ranges provided a clearer distinction between assessment categories and allowed for more precise quantification of results. This adaptation ensured greater consistency and objectivity in evaluation while enabling statistical analysis and data-driven decision-making. By combining descriptive feedback with these numerical ratings, the assessment captured both qualitative insights and measurable outcomes, offering a more comprehensive evaluation of the e-learning tools.

User feedback was collected through a researcher-designed feedback form, which was distributed after the students explored the CAA. The feedback form, focused on students' experiences with the CAA, was validated by five experts—two registered psychology professors who are also licensed psychometricians and three mathematics professors with Ph.D. degrees. While they are not from the fields of evaluation and educational technology, their combined expertise ensures a thorough validation process. The psychology professors, with their deep knowledge of assessment and psychometrics, examined the psychological aspects of the questions to ensure clarity, fairness, and alignment with sound measurement principles. Meanwhile, the mathematics professors, each with over 10 years of experience, have extensive exposure to assessment practices and are frequently consulted as validators for survey and interview instruments, proving their credibility in this role. This interdisciplinary approach enhanced the feedback form's clarity, reliability, and overall effectiveness, ensuring it accurately captured students' experiences and insights.

2.4 Data Analysis

The data analysis employed a mixed-methods approach, integrating both quantitative and qualitative techniques to comprehensively evaluate the CAA's content and functionality. Descriptive statistics, such as mean and standard deviation, were used to synthesize expert validation results, offering a clear summary of the system's performance. Simultaneously, user feedback underwent thematic analysis to identify recurring patterns and insights. Thematic analysis is a well-established qualitative research method that uncovers and interprets themes within a dataset, often providing deeper understanding and novel perspectives (Boyatzis, 1998; Elliott, 2018; Thomas, 2006). This process involves multiple readings of the data, with close attention to recurring ideas and patterns (Creswell & Creswell, 2018). In this study, thematic analysis was utilized by systematically examining user feedback to identify consistent themes related to the CAA's usability and accessibility, system design and features, learning experiences and skill development, as well as student feedback and suggestions for improvement. This method involved coding the data, grouping similar responses, and interpreting the significance of recurring ideas, providing a structured approach to understanding participants' perspectives. Thematic analysis is especially effective in capturing the meaning behind participants' language, enabling a detailed exploration of their experiences (Braun & Clarke, 2006; Liebenberg et al., 2020; Muhammad et al., 2023; Xu & Zammit, 2020). The insights gained from this analysis were crucial for highlighting strengths, pinpointing areas for improvement, and refining the CAA to better address students' needs and expectations.

3. Results and Discussion

3.1 Results

3.1.1 Needs Analysis Phase

In this study, the needs analysis involved a literature review, examining sources such as journals and conference proceedings related to NRP-solving, technology-enhanced assessments, and DDR. This phase began by recognizing gaps in assessing NRP-solving skills and the limitations of traditional assessment methods, highlighting the need for a computer-aided assessment (CAA). A review of relevant literature on NRP-solving, technology-enhanced assessments, and DDR methodology reinforced this necessity.

In the case of NRP-solving, the use of NRPs in training students was found to improve HOTS and problem-solving skills (Kablan & Uğur, 2019; Yazgan, 2015). According to Zulkipli and Saidah (2024), one approach to assessing students' problem-solving skills was to give them non-routine math problems, which require the use of HOTS. This approach, in turn, encouraged students to think more critically and deeply, especially when tests based on HOTS were administered (Yolida et al., 2023). Moreover, NRPs were also seen as a way to enhance communication and soft skills (Rusmini et al., 2021). This highlighted the importance of developing HOTS, which laid the foundation for defining the assessment content and goals. Polya (Legarde, 2022) suggested that NRPs should have been used to provide students with opportunities to develop HOTS through the processes of understanding, analysis, exploration, and the application of mathematical concepts. This deepened the understanding of NRP-solving competencies. Andrade et al. (2020) emphasized the importance of incorporating NRPs into mathematics curriculum materials, noting that this improved students' problem-solving performance. They suggested that students should have been introduced to various strategies so they could explore alternative approaches when one method failed. This approach made mathematics more engaging and enjoyable, allowing students to plan and create their own solutions. It also encouraged teachers to adopt a more flexible mindset, promoting diverse problem-solving methods rather than focusing on a single solution. Additionally, Kablan and Uğur (2019) highlighted that NRPs were situations that required HOTS and were rarely found in instructional materials. These findings served as empirical evidence supporting the need for specialized formative assessment tools in NRP-solving, which would act as a practice exercise test bank for students.

In light of this, one potential solution to enhance mathematical abilities was by engaging students in solving NRPs. However, preparing students with the necessary skills for this became more difficult, particularly due to the disruptions caused by the COVID-19 pandemic (Julie et al., 2017; Wijaya, 2020). This underscored the importance of technology-enhanced assessments. Bearman et al. (2023) emphasized that technologies were utilized in multiple ways to facilitate exam delivery, handle assignment submissions, enhance feedback on assessments, and, importantly, store and share assessment results. This reinforced the importance of technology in delivering assessments. E-assessment was found to help create innovative assessment practices that increased students' learning motivation and engagement (Marriott, 2009; Pham, 2022). This exploration of innovative e-assessment strategies underscored their impact on learning outcomes. Wang (2011) highlighted that the item formats used in e-assessments were limited to multiple-choice questions, true/false, matching, or short-answer responses, which might not have fully captured certain aspects of students' mathematical understanding, such as their logical and analytical abilities. This highlighted the need for computer-aided tools to evaluate complex skills. Tonbuloglu (2022) explored trends in e-assessment and highlighted how e-evaluation activities had evolved and transformed over time, influenced by advancements in technology, the pandemic, the growth of e-learning, increased communication opportunities, and various other factors. This reinforced the importance of adaptive and interactive assessment features.

Incorporating NRPs proved to be a valuable tool for enhancing HOTS. Since these problems were not easily solvable by routine methods, students had to engage in deeper cognitive processes. Working through NRP encouraged them to think critically, explore multiple solutions, and develop a better understanding of complex concepts. This significantly improved their ability to think critically and creatively, which are at the core of HOTS. While NRPs alone contributed to enhancing HOTS, integrating them with CAA provided additional benefits. CAA offered immediate feedback, adaptive learning pathways, and data-driven insights into student performance. This combination allowed for personalized learning experiences, where students could work on NRPs at their own pace, receive instant feedback on their thought processes, and adjust their strategies accordingly. This integration not only facilitated the development of HOTS but also provided students with the necessary tools to reflect on and improve their problem-solving strategies.

Finally, in the context of DDR, the DDR methodology was grounded in the work of Richey and Klein (2007) who outlined four key phases: analysis, design, development, and evaluation. Hasbullah (2022) highlighted the growing dominance of technology-based instruction in the design and development of product and tool research, further strengthening its role in educational innovation. DDR was utilized to create and refine interventions such as teaching and learning methods, educational materials, systems, educational gaming tools, and leadership models, all aimed at addressing complex educational challenges (Kragt & Day, 2020; Sahrir et al., 2012), thereby reinforcing the study's methodological approach. Additionally, Govindasamy et al. (2023) conducted a literature review illustrating the systematic approaches used in DDR, further expanding the theoretical foundation for the study's research design.

Collectively, these studies provided a strong framework for understanding the necessity of a computer-aided assessment for NRP-solving and underscored the importance of integrating technology-enhanced assessment methods into educational practice.

3.1.2 Design and Development Phase

3.1.2.1 Design Phase

The design phase established the foundation for the computer-aided assessment (CAA), defining its purpose, aligning it with course outcomes, determining assessment formats, selecting a platform, and choosing a name. The CAA aimed to expose students to solving non-routine problems (NRPs) by encouraging active engagement with these problems.

Mathematics in the Modern World (MMW) was identified as the most suitable course for integrating NRPs, focusing on two main topics: Mathematical Language and Symbols (sets, functions, relations, and logic) and Problem Solving and Reasoning (deductive and inductive reasoning, patterns, and recreational mathematics). To promote practice and feedback, the CAA used formative assessments with open-ended questions, requiring students to manually input their responses rather than relying on multiple-choice formats (Atlintas, 2022). Platform selection prioritized cost management and reliable services, balancing functionality with affordability. The CAA was named 'MatHOTSanayan,' combining 'Math,' 'HOTS,' and '*Sanayan*'—a Filipino term for consistent practice—which reflects the study's goal of exposing students to NRP-solving through regular engagement. This fusion of terms captured the CAA's focus on mathematics, higher-order thinking skills, and the importance of habitual learning and mastery.

The design phase built a solid foundation by aligning the CAA with the course outcomes of MMW while emphasizing active student engagement. The emphasis on formative assessment and open-ended questions reflected Altintas' (2022) assertion that open-ended items developed based on real-life situations provided more effective feedback in the formative assessment process. Likewise, Karakaya and Şata (2022, as cited in Altintas, 2022) emphasized that open-ended items fostered alternative ways of thinking in students, as they were deigned at the level of analysis, evaluation, and creation based on Bloom's taxonomy, as well at critical thinking, creative thinking, and problem-solving according to Haladyna's classification. Additionally, Kubiszyn and Borich (2003, as cited in Altintas, 2022) highlighted the importance of open-ended items in drawing inferences about complex higher-order cognitive skills, such as problem-solving, analysis, and evaluation. By centering the CAA around open-ended questions, students were encouraged to engage deeply with the material, formulate their own responses, and think critically—key components of HOTS.

International Journal of Didactic Mathematics in Distance Education
Volume 2, Issue 1, pp. 81-104, E-ISSN: 3047-9207
https://jurnal.ut.ac.id/index.php/ijdmde

Moreover, choosing MMW as the course for implementing the CAA was a deliberate decision due to its focus on mathematical language, symbols, reasoning, and problem-solving. These elements provided a strong foundation for developing open-ended questions, which played a key role in fostering meaningful learning experiences by encouraging students to reason, reflect, and articulate their understanding through language (Evans, 2020; Monrat et al., 2022).

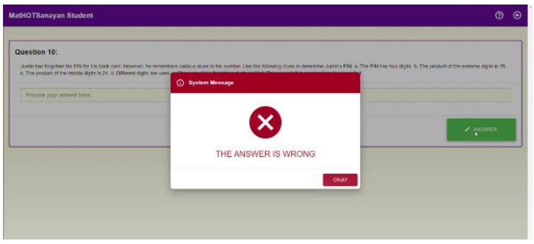
3.1.2.2 Development Phase

The development phase began with drafting the CAA's contents, consisting of NRPs. The CAA served as a test bank of formative assessments divided into four modules for the Mathematics in the Modern World (MMW) course: 30 NRPs each on mathematical language and symbols, deductive and inductive reasoning, patterns, and recreational problems using mathematics. After the first draft, the modules underwent expert validation. Five mathematics experts, all master's degree holders teaching MMW, reviewed the modules to ensure the problems and their intended purposes were accurately represented. Their feedback was incorporated to refine the CAA's contents. Once finalized, the NRPs and their answers were encoded into the system, allowing multiple correct answers and formats to capture diverse student responses. Specific instructions and an overview of each module were also written and encoded at the start of the modules.

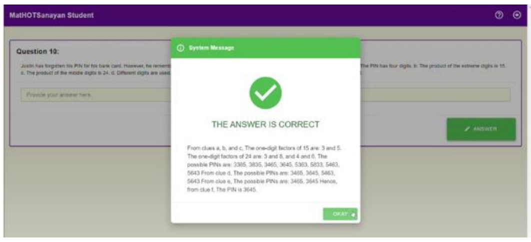
Some of the key features of the CAA include:

- *Immediate Verification Feedback*. Provided students with instant feedback on their answers, helping them learn from mistakes and stay engaged. Correct answers included sample solutions for reference in solving similar problems.

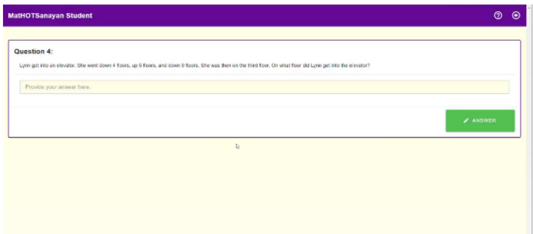
Immediate Verification Feedback for an incorrect answer:

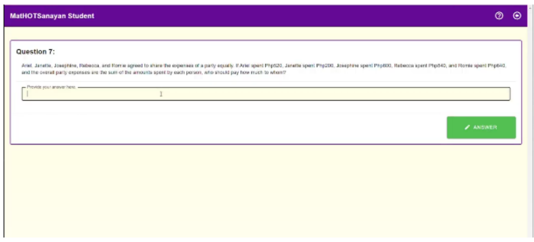


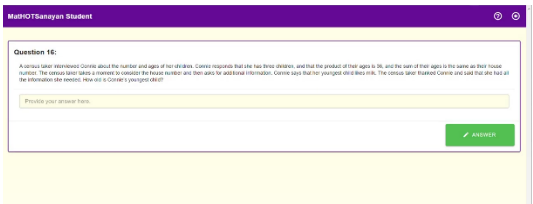
Immediate Verification Feedback for a correct answer with a sample solution

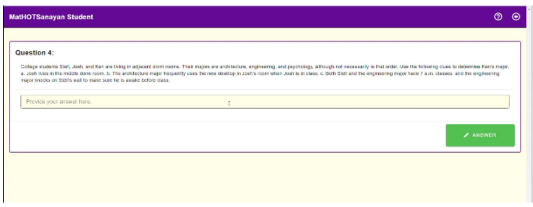


- *Focus on HOTS*. Encouraged students to solve complex problems, think critically, and apply their knowledge creatively.





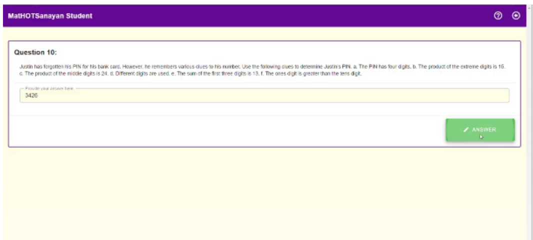




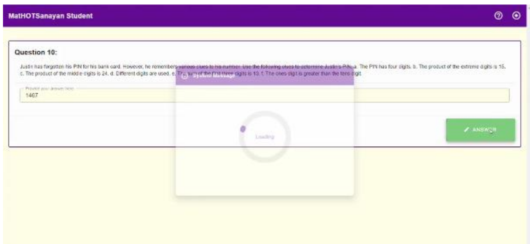
- *Triple Attempt Opportunities*. Allowed up to three attempts per problem to encourage persistence, resilience, and learning from errors, promoting a positive learning experience.

International Journal of Didactic Mathematics in Distance Education
Volume 2, Issue 1, pp. 81-104, E-ISSN: 3047-9207
https://jurnal.ut.ac.id/index.php/ijdmde

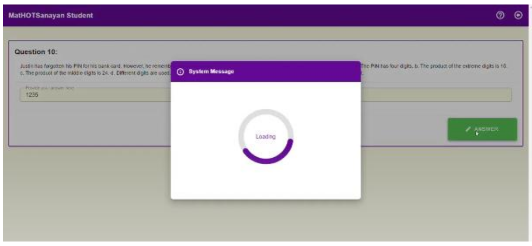
First Attempt:



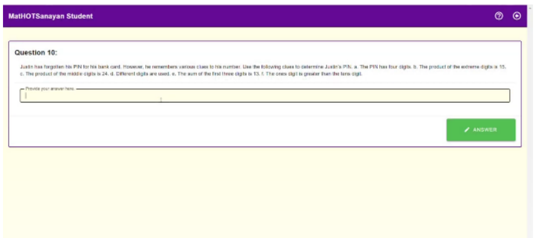
Second Attempt:

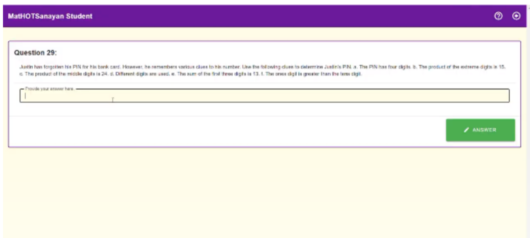


Third Attempt:



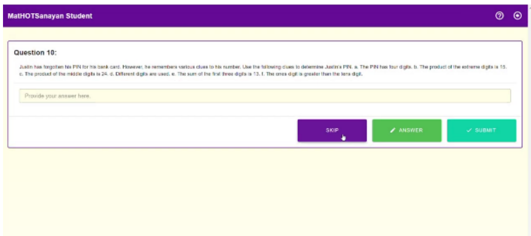
- *Randomized Problems Order*. Maintained student engagement by presenting the problems in different orders during each practice session. This approach prevented memorization and fostered a deeper understanding of the material.



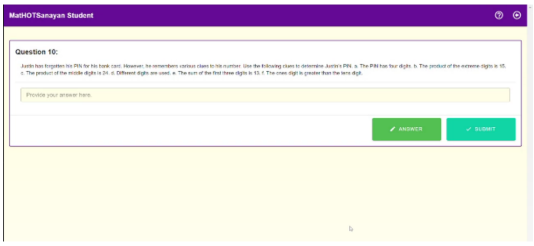


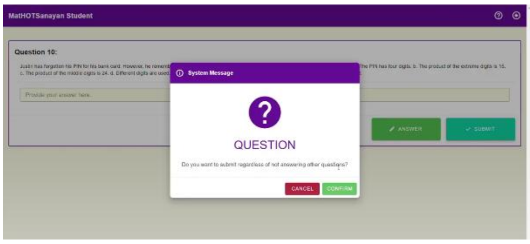
- *Skip and Revisit Options*. Allowed students to navigate the assessment flexibly, skipping a problem after three failed attempts and revisiting it later. This prevented exhaustion and let them manage their pace, focusing on areas of confidence.

Skip Option



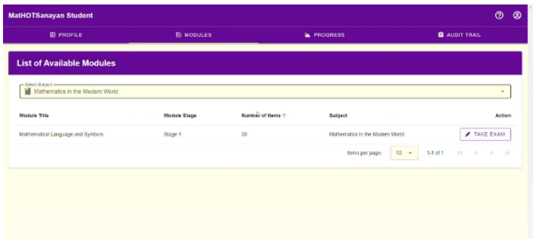
Revisit Options



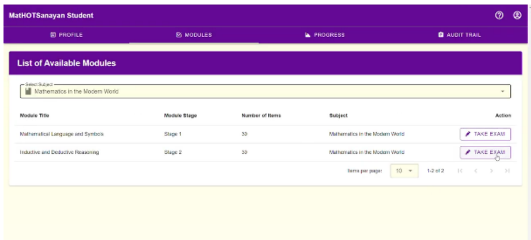


- *Sequential Module Unlocking.* Ensured that students mastered one topic before moving on to the next, promoting a solid foundation of knowledge.

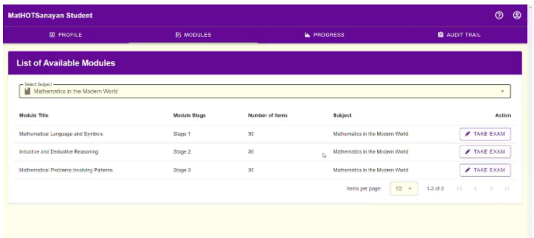
Only Module 1 is unlocked, while Modules 2 to 4 are still locked:



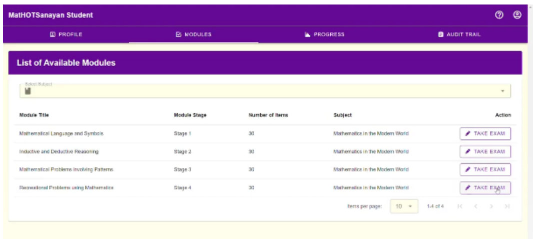
Modules 1 and 2 are unlocked, while Modules 3 and 4 remain locked:



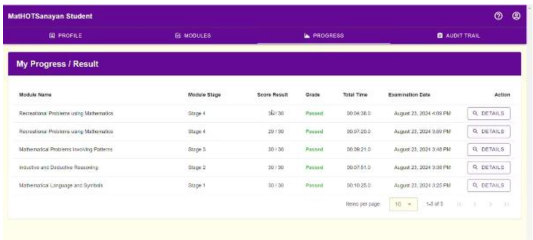
Modules 1 to 3 are unlocked, while Module 4 remains locked:



All Modules from 1 to 4 are unlocked:



- *Unlimited Access to Unlocked Modules.* Permitted continuous review and practice of completed modules to reinforce learning, helping them solidify their understanding and prepare for summative assessments at their own pace.



- *Automated Grading.* Provided students with immediate feedback on their scores, saving time for both students and teachers and allowing more focus on learning.

3.1.3 Evaluation Phase

This study adopted both approaches: quantitative methods were applied through expert validation using rubric-based evaluation results, while qualitative methods were used to analyze user experience through thematic analysis.

International Journal of Didactic Mathematics in Distance Education
Volume 2, Issue 1, pp. 81-104, E-ISSN: 3047-9207
https://jurnal.ut.ac.id/index.php/ijdmde

3.1.3.1 Expert Validation

Expert validators evaluated the CAA's usability, functionality, and alignment with the intended assessment objectives. Table 1 provides a summary of the expert validation results for different categories evaluating the CAA, showing their mean scores, standard deviations (SD), and verbal interpretations.

Table 1
Summary of Expert Validation Results for CAA

Category	Mean	SD	Verbal Interpretation
1. Functionality	2.50	0.00	Works Well
2. Accessibility	2.90	0.14	Works Well
3. Technical	2.85	0.14	Works Well
4. Mobile Design	3.00	0.00	Works Well
5. Privacy, Data Protection, and Rights	2.93	0.00	Works Well
6. Social Presence	1.60	0.15	Minor Concerns
7. Teaching Presence	2.20	0.18	Works Well
8. Cognitive Presence	3.00	0.00	Works Well

As shown in Table 1, the CAA was evaluated in terms of *Functionality*, *Accessibility*, *Technical*, *Mobile Design*, *Privacy*, *Data Protection and Rights*, *Social Presence*, *Teaching Presence*, and *Cognitive Presence*. The CAA's *Functionality* was evaluated in terms of Scale, Ease of Use, Tech Support/Help Availability and Hypermediality. As shown in Table 1, the results indicate that the experts found the CAA, in terms of *Functionality*, "works well", as indicated by an overall mean rating of 2.50 (S.D. = 0.00). This suggested that the experts believed the CAA served its intended purpose. The standard deviation of 0.00 indicated perfect agreement among the experts, which means all the experts rated it consistently. In terms of *Accessibility*, the CAA was evaluated based on accessibility standards, user-focused participation, required equipment, and cost of use. As depicted in Table 1, the results indicate that the overall mean rating for the CAA's Accessibility was 2.90 (S.D. = 0.14), which is interpreted as "works well." This suggested that the experts were convinced that the CAA was capable of supporting multiple learning approaches and fostering engagement for all students. The standard deviation of 0.14 showed that there was minor variation in the experts' opinions, which meant there was general agreement but slight differences in specific aspects of accessibility.

The CAA's *Technical* aspects was evaluated in terms of integrating/embedding within an LMS, desktop/laptop operating systems and browser, and additional downloads. An overall mean rating of 2.85 (S.D.= 0.14) was found for the CAA's *Technical* aspects, as shown in Table 1. This was interpreted as "works well", which means the experts were convinced the CAA had the basic technology needed to operate effectively. The standard deviation of 0.14 suggested minor variability in expert assessments, which meant while the experts were mostly in agreement, there were slight differences in their evaluations of the *Technical* aspects.

In terms of *Mobile Design*, the CAA was evaluated based on access, functionality, and offline access. However, the CAA did not support offline access, so this feature was rated as 'not applicable'. The overall mean rating for *Mobile Design* was computed by considering only the access and functionality features. As can be gleaned from Table 1, the results indicate that the overall mean rating for the CAA's Accessibility was 2.90 (S.D. = 0.14), which is interpreted as "works well." The results showed that the overall mean rating for *Mobile Design* was 3.00 (S.D. = 0.00), which is interpreted as "works well". This suggested that the experts unanimously agreed that the CAA could be effectively delivered using mobile devices. The perfect rating and standard deviation of 0.00 indicated complete consensus among the experts.

In terms of *Privacy, Data Protection, and Rights*, the CAA was evaluated based on Sign Up/Sign in, Data Privacy and Ownership, and Archiving, Saving, and Exporting Data. As shown in Table 1, the results showed that the overall mean rating for the CAA's Accessibility was 2.90 (S.D. = 0.14), which is interpreted as "works well." The results showed an overall mean rating of 2.93 (S.D. = 0.15) for *Privacy, Data Protection, and Rights*, which is interpreted as "works well". This suggested that the experts believed the CAA met important standards in privacy and data protection, which are essential for building user confidence and ensuring adherence to pertinent laws. The standard deviation of 0.15 suggested minor variation, which meant there was overall consistency in ratings, but some slight differences regarding *Privacy, Data Protection, and Rights*.

For the CAA's *Social Presence*, it was evaluated in terms of collaboration, user accountability, and diffusion. As can be gleaned from Table 1, the results indicate that the overall mean rating for the CAA's Accessibility was 2.90 (S.D. = 0.14), which is interpreted as "works well." The results showed that the overall mean rating for *Social Presence* was 1.60 (S.D. = 0.15), which was interpreted as "minor concerns". This indicated that the experts perceived potential shortcomings in the aspects that contribute to a safe and trusting environment necessary for effective collaboration and teamwork. While there were some positive aspects of social presence within the CAA, there were notable areas that required improvement. However, it was important to note that the CAA was intended for self-regulated learning, serving primarily as a test bank of formative assessments; thus, collaboration and teamwork were not applicable in this context. The standard deviation of 0.15 suggested minor variability in expert opinions, which meant that most experts shared similar concerns, but there were slight differences in how they viewed the importance of social presence in this particular setting.

In terms of *Teaching Presence*, the CAA was evaluated based on facilitation, customization, and learning analytics. As depicted in Table 1, the results indicate that the overall mean rating for the CAA's Accessibility was 2.90 (S.D. = 0.14), which is interpreted as "works well." The results showed an overall mean rating of 2.20 (S.D. = 0.18) for *Teaching Presence*, which was interpreted as "works well". This suggested that the experts believed the CAA supported teaching presence. Although the rating suggested that the experts' views were generally positive, there may still have been opportunities for further enhancements, particularly in the customization aspects, to optimize the teaching presence in the CAA. The standard deviation of 0.18 indicated slightly more variability in the experts' assessments compared to other categories, which meant there was a broader range of opinions regarding teaching presence and potential areas for improvement.

Finally, for the *Cognitive Presence*, the CAA was evaluated based on enhancement of cognitive task(s), higher-order thinking, and metacognitive engagement. As shown in Table 1, the results showed an overall mean rating of 3.00 (S.D. = 0.00) for *Cognitive Presence*, which was interpreted as "works well". The mean score of a perfect 3.00 reflected all experts' confidence in the CAA's ability to enhance cognitive tasks, HOTS, and metacognitive engagement. These elements aligned with the primary goal of the researcher in the conceptualization and development of the CAA. The standard deviation of 0.00 indicated complete agreement among the experts, which meant they unanimously believed the CAA effectively supported cognitive presence.

The findings of this study were similar to those of Pamungkas and Widodo (2019) in their evaluation of e-Learning Ecologies, where they also found that *functionality*, *accessibility*, *technical*, *mobile design*, and *cognitive presence* were all good or "works well". However, there were differing findings in the category *privacy, data protection, and rights*. In this study, these aspects worked well, whereas their study reported minor concerns. Additionally, there were differences in *social presence*; in this study, social presence has minor concerns, as it is important to note that the CAA in this study was intended for self-regulated learning, serving primarily as a test bank of formative assessments. Hence, making collaboration and teamwork less applicable in this context. In contrast, social presence worked well in their study. Moreover, the category of *teaching presence* was not included in their study, while it is reported as working well in this study.

3.1.3.2 User Feedback

As part of the evaluation phase, user feedback on the CAA 'MatHOTSanayan' was collected through validated feedback forms from 15 students. Their responses offered valuable insights into the CAA's usability, features, learning experience, and areas for improvement. Using thematic analysis, the feedback was organized into four main themes and 11 sub-themes, reflecting students' experiences, challenges, and perceptions of the CAA designed for NRP-solving.

Theme 1: Usability and Accessibility of the CAA

Sub-theme 1.1: Ease of Navigation and Interface Design

The usability and accessibility of an e-assessment platform heavily depended on its ease of navigation and the intuitiveness of its interface. The majority of the participants found 'MatHOTSanayan' user-friendly, citing its straightforward navigation and intuitive design. For instance, Participant 9 stated,

> *"MatHOTSanayan is user-friendly. The interface is simple and intuitive. You just need to select the module, begin with module 1, carefully read and understand the problems, attempt to solve them, and enter your answer. It's easy to determine if your answer is correct or incorrect as the website provides immediate feedback. Each module consists of only 30 questions, specific to the module you are working on. Additionally, there is no requirement to show your solution, so any method you use to arrive at the correct answer is acceptable. With four modules covering different areas of mathematics, MatHOTSanayan offers a diverse learning experience."*

This comment underscored the CAA's simplicity, immediate feedback, and flexibility in problem-solving.

> Participant 12 explained, *"I think overall it is user-friendly, it was easy to use and learn. All of the needed tools are presented clearly and all of the tools you might need are arranged in a manner that you won't have to think twice about where to click next. …"*

These responses supported the findings of Almuhanna et al. (2024), who found that students affirmed the significance of having an easy-to-use e-assessment.

Sub-theme 1.2: Accessibility and Compatibility

Accessibility and compatibility across devices and settings were crucial for ensuring a seamless learning experience. Participants consistently emphasized the convenience of using 'MatHOTSanayan' on different devices. For instance, Participant 6 shared.

> *"It is user-friendly; it is easy to navigate and highly accessible using a cellphone and laptop, from selecting a subject to taking the exam. The platform loads swiftly, too. …"*

This highlighted the CAA's compatibility with multiple devices, clear layout, and smooth user experience. Similarly, Participants 12 and 14 highlighted MatHOTSanyan's responsive and user-friendly design, noting its seamless usability and adaptability across desktop and mobile devices.

The participants' responses confirmed one of the benefits listed by Sikurajapathi et al. (2020), which was that students could access it at different times and from different locations. Similarly, Almuhanna et al. (2024) found that students viewed the flexibility and convenience of taking the exam anytime and from any location as one of the most significant advantages of computer-based exams.

Sub-theme 1.3: Technical Challenges

Despite the overall positive feedback, participants reported several technical challenges that affected their user experience. These issues included occasional glitches, slow loading times, and problems with the answer input format. For instance, Participant 5 expressed frustration.

> *"MatHOTSanayan was buggy when it initially released. The modules wouldn't appear when they did, I couldn't even submit them for a while. Sometimes, the images wouldn't even load, so I would have to abandon the attempt because there is no way for me to answer the question unless I've already encountered it before and knew the answer. This would often happen when I am around 25 questions into the attempt, which makes another attempt quite tedious and annoying. …"*

Similarly, other participants, like Participants 1, 3, and 6, reported experiencing technical issues such as missing or incorrect images, bugs, and glitches, often requiring page refreshes or starting over — leading to frustration and suggesting the need for system improvements. These technical issues suggested that the platform's functionality needed to be refined in order to improve user experience. This was in line with findings from a study by Gunesekera et al. (2019), which highlighted that usability issues in e-learning systems affected user satisfaction and relationships, resulting in a lack of user continuity.

Several participants also noted that even when their answers were correct, minor formatting differences often led to incorrect submissions. For instance, Participant 2 shared: *"There were challenges such as there are formats of answers, I just hope that the formats are even more broad or there are many* *so that our answers would get in right away. Sometimes even though I followed the format it still doesn't work."* While Participant 4 described as frustrating experience: *"Despite knowing the right answer, most of the questions in the last chapter would not accept the answer typed in even if it is essentially the same thing. ..."*

Participants 5, 11, 12, and 14 similarly expressed dissatisfaction with the platform's answer recognition system, highlighting issues like the lack of flexibility in accepted answer formats, the difficulty of inputting multiple answers in one tab, and the system's sensitivity to minor variations like spaces or symbols, leading to confusion, frustration, and discouragement among students.

These issues with the input format highlighted the need for improvement in user experience design. While the interface was generally perceived as accessible and straightforward, the precision required in the answer box negatively impacted the overall usability. This observation aligned with research highlighting the importance of user-friendly design in educational tools. Alqurni (2023) stressed that the efficacy of an e-learning system heavily depended on how well-designed its user interface was and how easy it was for students to use. The results also highlighted how a well-designed user interface guaranteed that students could easily navigate the platform, keeping them focused on learning rather than troubleshooting the interface.

Theme 2: Features and System Design

Sub-theme 2.1: Sequential Module Unlocking and Unlimited Access to Unlocked Modules

Students expressed mixed feelings about the system of sequential module unlocking — while some appreciated its structured approach, others felt restricted by it. On the other hand, unlimited access to unlocked modules was widely seen as a strong advantage. This balance between structure and flexibility shaped participants' overall experience with the platform.

The majority of participants indicated that completing modules in a predetermined sequence helped them maintain organization and focus, mastering one concept before moving on to the next. For instance, Participant 1 explained:

> *"I think that it did help me stay organized as I was able to focus all my attention on mastering one concept instead of dividing it into various concepts that have different content. Additionally, it helped me build foundational knowledge as the concepts or the modules are arranged in a way that the first module could be beneficial in answering the next one, and so on."*

Participant 2 echoed this sentiment and noted:

> *"Completing and submitting them in a predetermined sequence was nice because the more that I engage with the modules consecutively, I learn more since I started from the bottom and to the top. Learning really progressed. It helped me to stay organized because as someone who had a shorter attention span and indecisiveness, it gave me no other choice and it was more challenging (in a nice way) in a way that I know I need to finish this one in order to proceed."*

Similarly, Participants 3 and 6 emphasized how the step-by-step approach allowed them to focus on one concept at a time, making it easier to build foundational knowledge and prepare for subsequent topics.

Despite the perceived advantages, some participants encountered challenges, particularly when they struggled with specific modules. Participant 10, for instance, expressed frustration: *"Having the modules locked did not help me at all since I was stuck at the first module. ..."*

This issue was also raised by Participant 11, who shared:

> *"... But I think I am speaking for a lot of students when I say that how it was made was stressful. Since we were in a time limit crunch with a designated deadline (for us, it was determined to be a module a day for four(4) consecutive days), it was stressful because it takes a really long time to answer the module, around 1 and a half hours or 2. When you fail the subject, you need to keep repeating and repeating the module when we have other responsibilities to do in life..."*

Sub-theme 2.2: Self-Paced Learning and Autonomy

Participants highlighted the flexibility and convenience of personalized learning offered by the platform compared to traditional methods. This self-paced approach allowed students to focus on their individual learning needs and improve their understanding at their own speed. For instance, Participant 9 said: "*... MatHOTSanayan allows me to learn at my own pace and focus on areas where I need improvement, providing a more personalized and effective learning experience compared to traditional classroom settings."*

Participant 6 added:

> *"Compared to pen-and-paper assessments, using the MatHOTSanayan allows you to answer at your own pace and even explains how to solve a problem. In the traditional learning method, the solutions will enter and exit through the student's ear, and they might not remember it; however, with MatHOTSanayan, you can copy it online. Of course, unlike the traditional assessment, you can input your answer and eventually know if it's correct."*

Meanwhile, Participants 14 and 2 offered valuable insights into the platform's convenience and personal impact, with Participant 14 focusing on how it could have improved their learning during the pandemic and Participant 2 reflecting on the confidence and HOTS they had gained from using the platform.

These findings supported those of Kumari et al. (2021), who came to the conclusion that online delivery and scheduling were more flexible.

Theme 3: Learning Experience and Skill Development

Sub-theme 3.1: Support for Non-Routine Problem Solving and HOTS

The platform had played a significant role in enhancing participants' critical thinking and problem-solving abilities. Many participants shared how their exposure to NRPs challenged them to apply their knowledge in diverse and creative ways. For instance, Participant 1 mentioned:

> *"It definitely improved my Higher-order thinking skills as it exposed me to NonRoutine Problems that require critical thinking. This challenged me to apply the methods or knowledge that I have in various ways. It helped me deconstruct the problem into multiple pieces of information before getting the correct answer to the whole problem. It helped me develop my skills in organizing information."*

Participant 9 added:

> *"I believe MatHOTSanayan has improved my Higher-Order Thinking Skills (HOTS) by challenging me to think critically and creatively when solving mathematical problems. Rather than relying on rote memorization or simple calculations, the program encourages me to analyze problems from different perspectives, apply logical reasoning, and develop innovative strategies to arrive at solutions. This not only strengthens my mathematical abilities but also enhances my overall cognitive skills."*

Several participants emphasized that the platform enhanced their problem-solving and critical thinking skills. For instance, Participant 1 clearly conveyed their experience, stating: *"All I want to say is that I had a positive experience with MatHOTSanayan. It really helped me improve my problem-solving skills and critical-thinking abilities. ..."*

Other participants also shared similar thoughts. Participant 3 highlighted how MatHOTSanayan positively influenced their learning, particularly in solving NRP problems. Participant 6 emphasized how the platform broadened their understanding of various non-routine problems and ways to solve them. Finally, Participant 9 reflected on how MatHOTSanayan rekindled their interest in solving math problems, improving their problem-solving skills and overall mathematical abilities.

Several participants emphasized the platform's support for creative thinking and reflection. For instance, Participant 6 shared:

> *"Indeed, there have been instances where I became upset when I made a mistake, but I later realized that I had misinterpreted the question. As a result, I carefully analyzed the problem, and my reading comprehension improved. Most importantly, my creativity has been enhanced because you have to look outside the box and picture yourself involved in the problem. My problem-solving ability was further developed by doing numerous trial-and-*

> *error problems. Finally, it helped with my memory since I could retake the test as many times as I wanted, so I knew the answer once I knew the question."*

Participant 15 expressed their appreciation for the freedom the program provided in approaching questions:

> *"By not giving me an instruction how to solve it like it allows me to understand each question in my own understanding and came up with an answer based on what I really thought is right. It's not like computing or solving an equation where you apply specified formula to get the answer it more like how I understand the question, so I really go beyond the concepts I knew in the subject but rather being able to realize it's application in different approaches. Additionally, giving me freedom to choose, like in completing the sentence where 2 are unknowns, I am the one to decide which is which that I think will complete the sentence. This questions really made me think what the right pair of all the possibilities is because multiple pair of answers can be made but only one is actually correct that satisfy the sentences once completed."*

Other participants echoed these sentiments. Participant 5 described how the program improved their HOTS and encouraged them to find more efficient or creative solutions rather than relying on repetitive approaches. Similarly, Participant 13 reflected on how the program pushed their boundaries in problem-solving, enabling them to think more clearly and develop efficient strategies through pattern recognition and rule-setting.

These responses confirmed Pitorini et al (2024) findings that a PBL-based e-module combined with Socratic dialogue had effectively enhanced students' critical thinking skills. Students who had utilized the e-module demonstrated greater improvement in critical thinking than those who had not. Similarly, Elbyaly and Elfeky (2023) discovered that the experimental group's development of critical thinking skills had been positively and significantly impacted by a problem-solving program run on the Blackboard platform.

Sub-theme 3.2: Quality of Feedback and Guidance

The quality and immediacy of feedback provided by the platform played a crucial role in the participants' learning process. Several participants highlighted the benefits of receiving immediate feedback. For instance, Participant 14 shared: *"Immediate feedback is beneficial for students who are self-studying, especially in remote learning setups. It also provides iterative practice opportunities for the questions I got wrong."*

Participant 2 added: *"After you get the right answer MatHOTSanayan would show a guide strategy on how they come up with that answer and for me it enlightens the students like me more."*

Participant 15 elaborated on how feedback improved their learning:

> *"By making me answer the question in my own way. I really solved every problem based on what I think is applicable concept to it and solve it in my own approach. Then after answering it correctly, the system will give a message on how it was solved which most of the time is different in my own way of solving. With that, I able to learn different techniques that is helpful in answering other questions where my own technique did not work for some reason. Moreover, being able to know other techniques, I understand the concept more and able to combine it with my own approach."*

These findings supported Yu and Cai's (2022) research on the importance of instant feedback for the effectiveness of online learning. Their study demonstrated that immediate feedback was essential for encouraging students to continue using the program and enhancing their engagement.

Sub-theme 3.3: Flexibility and Control

The platform's features provided students with greater control and flexibility over their learning experience. By offering multiple attempts, the ability to skip and revisit questions, and a structured sequence of modules, the platform empowered students to take ownership of their learning journey. These features fostered persistence, experimentation, and self-paced learning, aligning with effective educational practices that prioritize adaptability and deeper engagement.

Many participants appreciated the opportunity to learn at their own pace and reflect on their learning strategies. For instance, Participant 7 shared their strategy for using the platform's features:

> *"I figured that there were two types of attempts I made while using the platform. My initial attempt for each module was experimental, serving as an approach to grasp the problems through trial and error and using a solution I hastily made. The next attempt involves revising my solutions to make them more effective. Thus, I found the reattempt feature for both the items and the module itself quite useful."*

Participants 2 and 12 also expressed appreciation for the platform's supportive features. The general sentiment indicated that the platform's flexible and interactive approach positively impacted their learning outcomes by encouraging experimentation, reflection, and consistent goal-setting.

Several participants highlighted the value of the three-attempt feature in encouraging reflection and deeper understanding. For instance, Participant 1 explained,

> *"I think that the feature of offering three attempts to solve each problem before moving on to the next problem actually encourages me to persist in solving the problem even if I am having a difficult time solving it. Additionally, for every attempt that I made, it led me to reflect on my answer and try to think of the mistakes or errors that I made which might result in a deeper understanding of the problem. It also allowed me to try different strategies in answering the problem."*

This sense of persistence and reflection was echoed by other participants. Participant 2 noted that having multiple attempts made them more eager to solve the problem, while Participant 9 highlighted how it enhanced their learning and encouraged them to explore different methods. Both found that with each try, their understanding deepened, and their strategies improved.

These responses supported Wang et al. (2020), who argued that counting the number of times students tried to answer a question correctly was a good way to gauge persistence propensity — the tendency to keep trying despite difficulties. In essence, students were more persistent the more attempts they made.

A few participants also reported a reduction in anxiety associated with problem-solving due to the skip and revisit feature. For instance, Participant 8 stressed:

> *"This feature helped me get higher scores since it allowed me to redeem myself from skipped questions with the additional knowledge I got from other questions. ..."*

Participant 13 added: *"... It makes me think of my mistakes, and answering correctly other questions before revisiting them gives me a sort of confidence boost."*

Similarly, Participant 14 mentioned:

> *"There is a degree of learning reinforcements through revisiting those questions that I initially skipped. On my end, I feel satisfied having given the chance to answer those questions again. It gave me more time to think about it, as well as a refreshed perspective on how to tackle the question."*

Moreover, many participants expressed that the predetermined sequence of modules helped reduce feelings of overwhelm and stress. For instance, Participant 13 stated:

> "*... It helped me not be overwhelmed by the amount of work, and take all the time that I need for each module. If all of the 4 modules was set to show I wouldn't focus on mastering each one rather I would focus on completing it. Giving me a short and attainable goal helped me take that task one step at a time rather than feeling like a daunting big leap I have to take.*"

Similarly, Participants 12 and 15 expressed how focusing on one module at a time helped them stay organized, avoid feeling overwhelmed, and take the time needed to master each topic.

The findings reflected the "spacing effect," or "distributed practice," which described how gaps in study sessions could enhance later memory retention (Emeny et al., 2021). Psychologists emphasized that students were better able to remember information when it was presented across multiple sessions spaced apart in time (Ellwood, 2021). Similarly, Emeny et al. (2021) discovered that students who engaged in spaced practice outperformed those who practiced in large blocks on math assessments.

Theme 4: Student Feedback and Suggestions for Improvement
Sub-theme 4.1: Constructive Feedback for Platform Enhancement

Participants provided constructive feedback and valuable suggestions for enhancing the MatHOTSanayan platform. Their responses highlighted areas for improvement related to user-friendliness, technical issues, interactive features, and answer formats.

A key area of improvement identified by participants was the platform's user-friendliness. For instance, Participant 8 mentioned:

> *"I think MatHOTSanayan has a lot of potential to be an effective website that focuses on sharpening people's higher-order thinking skills; it just needs some improvements to make it more user-friendly. …"*

Other participants shared similar thoughts. Participant 15 pointed out issues with how questions and answer choices are displayed, suggesting clearer formatting for better readability. Participant 11 emphasized the website's potential and encouraged refining it to better support students' learning, offering constructive feedback for improvement.

Many participants also emphasized the need to address technical issues. For instance, Participant 1 suggested: *"… I do not have any particular suggestion for improvement except to fix those bugs."*

Participant 8 echoed this sentiment: *"…For suggestions, I wish this technical difficulty be fixed."*

In addition to technical improvements, participants recommended incorporating interactive features to enhance engagement. For instance, Participant 7 recommended:

" … I recommend adding a component below these problems, such as fillable boxes, to ensure the correct format is maintained: gave an amount of Php. Another way, is to have item-specific instructions. I do think this would require a lot of work since there are many items, but maybe it can only be applied to problems that have longer answers. …"

Several participants requested an auto-save functionality to prevent loss of progress. For instance, Participant 8 suggested: *"… For suggestions, I wish there is an auto save feature just like in MS Forms and Google Forms. …"* Similarly, Participant 12 recommended a save-progress feature to prevent frustration from lost work, emphasizing how this would improve the user experience.

Participants also called for improvements in the answer format and feedback system. For instance, Participant 8 suggested: *"… For suggestions, I wish there would be a provided example of the answer format for students to follow. …"* Similarly, Participant 15 proposed that when identical answers are input multiple times, the system could offer hints or clarifications to improve communication and help users understand what might be going wrong.

A few participants suggested adding hints and feedback after incorrect answers. For instance, Participant 10 added:

> *"… It would be better if the platform guided us through our wrong answers instead of gatekeeping the information and allowing us to read the explanation right after we get it right."*

However, this suggestion contradicted the purpose of the CAA, which was a test bank of formative assessments designed for non-routine problem solving. Providing hints in this context did not apply, as the focus was on encouraging students to engage with complex, unfamiliar problems rather than relying on immediate guidance.

Others echoed this sentiment, like Participant 6, who hoped for accessible solutions after answering, and Participant 13, who proposed a separate tab with answers as hints for tackling similar problems.

This feedback reflects a key limitation of the CAA 'MatHOTSanayan,' which incorporated formative assessments and was designed using this process. It did not offer conceptual support prior to the test, limiting its application mainly to assessment rather than instructional purposes.

Sub-theme 4.2. Likelihood to Recommend

Many participants expressed appreciation for the interactive nature of the platform and its potential for enhancing learning. For instance, Participant 1 shared:

> *"I would recommend MatHOTSanayan to a friend because it is actually a fun and interactive platform that could enhance one's adaptive learning capabilities. The aspect that I find particularly noteworthy is the feedback feature as one could possibly answer the problem with*

a different solution than the one in the feedback exposing them to an alternative way of thinking in problem-solving."

Similarly, Participant 7 appreciated the answer key feature, which provided immediate access to the "preferred" solution and helped them refine their problem-solving approach by identifying more efficient strategies.

Some participants found the multiple-attempt feature and progress tracking particularly beneficial. For instance, Participant 3 shared:

"I would recommend it to a friend because I know that this can significantly help them due to its features. The ability to attempt problems more than one which allows more understanding of concepts. The option to revisit questions and the sequential learning of MatHOTSanayan is the commendable aspects."

Similarly, Participant 4 noted:

"I would, besides reviewing, it can also be helpful in passing time. The graphs by the end of the quiz are amusing to look at since I could see my progress and how much I improved in between attempts."

These findings aligned with the study of Hughes et al. (2020), which emphasized that multiple attempts at the "weekly participation task" gave students who would have otherwise struggled with academic assessments a chance to succeed.

One notable benefit mentioned by participants was the absence of time limits. Participant 12 stated: For instance, Participant 12:

"... One of its standout features is the absence of a time limit. It allows me to take time to understand and grasp the problem without feeling rushed. It reduces stress and pressure which fosters a more relaxed and enjoyable learning environment."

This response supported the findings of Zakaria et al. (2022), who found that students should have been given extra time to finish tests and that time allocation should not have been overly strict in order to address some unavoidable problems.

Despite these positive aspects, some participants were hesitant to fully endorse the platform due to its perceived lack of conceptual support and usability issues in its early development. For instance, Participant 5 remarked:

"... I would only offer them the program if only for evaluation purposes alone because MatHOTSanayan doesn't have any modules to help the user understand fundamental concepts and processes before they begin taking tests. While we were provided knowledge through our MMW classes, MatHOTSanayan cannot stand on its own if it was released as is. Though, it is beneficial because if the user doesn't have as great of a higher-order thinking skill than he thought he would have, he could improve it by taking the exam multiple times. ..."

This response contrasted with the definition of e-assessment provided by Kundu and Bej (2021), who defined e-assessment as the process of creating, distributing, grading, and providing students with feedback on diagnostic, summative, and formative assessments using a computer and/or the internet. The CAA 'MatHOTSanayan,' designed for formative assessments, adhered to this process but was limited in its application, focusing primarily on assessment rather than instructional purposes.

Participants also pointed out usability and design limitations as barriers to recommending the platform. Participant 10, for instance, stated:

"I would not recommend it to others given its early development. Changes must be made for it to be user-friendly and be a platform that is a helpful tool to learn math."

Participant 11 added: *"If I will be completely honest, in its current state, I will not yet recommend it to a friend."*

These comments reflected participants' understandable concerns and hesitations about recommending the platform due to its current usability and design limitations. These responses aligned with the findings of Zakaria et al. (2022), who reported that while students responded favorably to online assessments, there were still many areas in need of improvement.

Limitations

Despite its promising outcomes, this study had several limitations. First, the sample size for user feedback was relatively small, involving only 15 students from three educational institutions, which may limit the generalizability of the findings. Second, the study focused on non-routine problem-solving in the context of Mathematics in the Modern World (MMW), and its applicability to other subjects remains unexplored. Third, while expert validation provided valuable insights into content quality, usability, and alignment with instructional goals, the study did not include long-term assessments of student performance and retention. Fourth, the effectiveness of the CAA in enhancing Higher-Order Thinking Skills (HOTS) and NRP-solving was not evaluated. It should be noted that the quality of students' HOTS could not be properly assessed, particularly because their familiarity with using computers was limited. This lack of familiarity may have further hindered their ability to effectively engage with the CAA, potentially affecting their HOTS development. Lastly, minor technical challenges were noted, indicating the need for further technical refinements to enhance system stability and user experience. Future studies could address these limitations by expanding the sample size, exploring cross-disciplinary applications, conducting longitudinal evaluations, and specifically assessing the impact of the CAA on the enhancement of HOTS and NRP-solving skills through controlled studies and long-term tracking. Additionally, future research should consider evaluating the quality of students' HOTS more thoroughly, ensuring that any limitations in students' computer familiarity are addressed through targeted training or support, which could enhance their ability to engage with the CAA and, in turn, foster the development of HOTS more effectively.

4. Conclusion

The needs analysis phase served as a critical foundation for identifying gaps and limitations in traditional assessment methods and the potential of computer-aided assessment (CAA) to address these issues. Through a comprehensive literature review, the study highlighted the importance of exposing students to non-routine problems (NRPs) and the role of technology-enhanced assessments in fostering higher-order thinking skills (HOTS), engagement, and performance. This analysis reinforced the need for a specialized, innovative assessment approach that leverages technological advancements to overcome challenges faced in contemporary educational practices.

The design and development phase of the CAA established a comprehensive and innovative foundation, ensuring alignment with the course outcomes of Mathematics in the Modern World (MMW) and fostering students' HOTS through engagement with NRPs. In the design phase, NRPs were integrated into key topics like mathematical language and symbols, and problem-solving and reasoning, creating opportunities for students to develop critical and creative thinking, as well as problem-solving abilities. The emphasis on formative assessment and open-ended questions supported meaningful learning experiences, encouraging students to articulate their understanding and explore multiple solutions. Careful consideration in selecting the platform and the name of the CAA—called 'MatHOTSanayan'—reflected the study's purpose of promoting consistent practice and deep cognitive engagement. The development phase involved a careful and deliberate process, resulting in a thoughtfully designed assessment tool. The CAA's four modules covered essential mathematical topics with a diverse range of NRPs, and its key features—such as immediate verification feedback, automated grading, triple attempts opportunities, and skip and revisit options—promoted active learning, resilience, and mastery of concepts. By integrating HOTS, randomized problems, and unlimited access to unlocked modules, the CAA ensured continuous practice and meaningful engagement, supporting individualized learning and preparing students to apply mathematical principles in creative and critical ways.

The evaluation phase underscored the CAA's reliability as a digital assessment tool and a well-validated, supportive educational resource, demonstrating its ability to support self-regulated learning through a well-structured, user-friendly interface. Through expert validation and user feedback, the CAA demonstrated commendable functionality, accessibility, and support for HOTS, particularly excelling in mobile design and cognitive presence. Students' positive experiences with the intuitive interface, cross-device compatibility, and immediate feedback further underscored the CAA's potential to foster engagement and individual growth. While some technical challenges and minor design limitations were noted, these issues presented opportunities for future enhancements without undermining the platform's overall stability and quality. With these improvements, the CAA holds significant potential for broader adoption in online education, contributing to a more engaging, interactive, and accessible online learning environment and student-centered learning experiences.

In light of these findings, future studies should assess the CAA's long-term impact on students' higher-order thinking skills and performance, its adaptability across disciplines, and its strengths compared to other tools. Exploring adaptive learning and collaborative features could further enhance personalized and interactive learning, refining the CAA as a transformative educational tool. Through these efforts, the continued development of the CAA can maximize its potential to support meaningful, student-centered learning and prepare learners for complex, real-world challenges.

Author Contribution

Author 1,2: Conceptualization, Writing - Original Draft, Editing and Visualization; Writing - Final Draft, Review & Editing, Formal analysis, and Methodology
Author 3: Validation and Supervision

Conflict of Interest

The authors declare no conflict of interest.

5. References

Ahmed, S.K. (2024). How to choose sampling technique and determine sample size for research: A simplified guide for researchers. *Oral Oncology Reports, 12:100662.* https://doi.org/ 10.1016/j.oor.2024.100662

Alqurni, J. S. (2023). Evaluating the user interface and usability approaches for e-learning systems. *International Journal of Information Technology and Web Engineering, 18(1), 1-25.* https://doi.org/10.4018/IJITWE.333638

Almuhanna, M. A. (2024). Teachers' perspectives of integrating AI-powered technologies in K-12 education for creating customized learning materials and resources. *Education and Information Technologies*, 1-29. https://doi.org/10.1007/s10639-024-13257-y

Andrade, R.R., Fortes, E.C., & Mabilangan, R.A. (2020). Problem solving heuristics and mathematical abilities of heterogeneous learners. *Universal Journal of Educational Research 8(11), 5114-5126.* https://doi.org/10.13189/ujer.2020.081111

Anstey, L., & Watson, G. (2018). Rubric for elearning tool evaluation. *Centre for Teaching and Learning Western University.* https://teaching.uwo.ca/pdf/elearning/Rubric-for-eLearning-Tool-Evaluation.pdf

Atlintas, O. (2022). The Use of open-ended items for giving feedback during the formative assessment process. *International Journal of Assessment Tools in Education 9, Special Issue, 109-125.* https://doi.org/10.21449/ijate.1132981

Barana, A., Boetti, G. & Marchisio, M. (2022). Self-assessment in the development of mathematical problem-solving skills. *Education Sciences, 2022, 12, 81.* https://doi.org/10.3390/edusci12020081

Bearman, M., Nieminen, J.H., & Ajjawi, R. (2023). Designing assessment in a digital world: An organising framework. *Assessment & Evaluation in Higher Education, 48(3), 291-304.* https://doi.org/10.1080/02602938.2022.2069674

Berisha, V. (2015). The general characteristics of mathematics textbooks for lower secondary school in Kosovo. *International Journal of Novel Research in Education and Learning, 292), 19-23.* https://www.noveltyjournals.com

Bianchi, N., Song, H., & Lu, Y. (2022). The effect of computer-assisted learning on students' long-term development. *Journal of Development Economics, 158*, 102919. https://doi.org/10.1016/j. jdeveco.2022.102919

Boyatzis, R.E. (1998). Transforming qualitative information: The matic analysis and code development. Sage Publications, Inc. https://psycnet.apa.org/record/1998-08155-000

Braun, V., & Clarke, V. (2006). Using thematic analysis in psy chology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Creswell, J.W., & Creswell, J.D. (2018). Research design: qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE Publications, Inc. https://spada.uns.ac.id/pluginfile.php/510378/mod_resource/content/1/creswell.pdf

Davies, B., Smart, T., Geraniou, E. & Crisan, C. (2022). STACKification: automating assessments in tertiary mathematics. *Twelfth Congress of the European Society for Research in Mathematics Education (CERME12), Feb 2022, Bozen-Bolzano, Italy. Hal-03750584*. https://hal.science/hal-03750584v1/document

Dyrvold, A., & Bergvall, I. (2023). Computer-based assessment in mathematics: Issues about validity. *LUMAT: International Journal on Math, Science and Technology Education, 11*(3), 49-76. https://doi.org/10.31129/LUMAT.11.3.1877

Elbyaly, M.Y.H. & Elfeky, A.I.M. (2023). The impact of problem-solving programs in developing critical thinking skills. *European Chemical Bulletin, 12(6), 6636-6642*. https://doi.org/10.31838/ecb/2023.12.si6.588

Elliott, V. (2018). Thinking about the coding process in qualitative data analysis. Qualitative Report, 23(11), 2850–2861. https://doi.org/10.46743/2160-3715/2018.3560

Ellwood, B. (2021). Spacing math practice across multiple sessions improves students' test scores and helps them accurately judge their learning. *PsyPost*.

Emeny, W. G., Hartwig, M. K., & Rohrer, D. (2021). Spaced mathematics practice improves test scores and reduces overconfidence. *Applied Cognitive Psychology*. https://doi.org/10.1002/acp.3814

Evans, C. M. (2020). Measuring student success skills: A review of the literature on collaboration. Dover, NH: National Center for the Improvement of Educational Assessment. https://files.eric.ed.gov/fulltext/ED607774.pdf

Fan, L. & Zhu, Y. (2000). Problem solving in singaporean secondary mathematics textbooks. *The Mathematics Educator, 5(1), 117-141*. https://repository.nie.edu.sg/server/api/core/bitstreams/ 978a61aa-a672-4f1a-ba95-605fdb26d399/content

Govindasamy, P., Abdullah, N., & Ibrahim, R. (2023). Employing modified design and development research (DDR) approach to develop creativity clay module to teach special educational needs (SEN) students with learning disabilities. *International Journal of Academic Research in Business and Social Sciences, 13(3), 461 – 476*. https://doi.org/10.6007/IJARBSS/v13-i3/15773

Gunesekera, A.I., Bao, Y. & Kibelloh, M. (2019). The role of usability on e-learning user interactions and satisfaction: a literature review. *Journal of Systems and Information Technology, 21(3), 368-394*. https://doi.org/10.1108/JSIT-02-2019-0024

Hasbullah, N.A., Fabil, N., Yusoff, A.M., & Azmi, A.S. (2022). Adapting a design and development research (DDR) approach in designing the mobile application model for tarannum subject. *Advances in Humanities and Contemporary Studies 3(2), 193-204*. https://doi.org/10.30880/ahcs.2022.03.02.018

Hughes, M., Salamonson, Y., & Metcalfe, L. (2020). Student engagement using multiple-attempt "Weekly Participation Task" quizzes with undergraduate nursing students. *Nurse Education in Practice, 46, 1-6*. https://doi.org/10.1016/j.nepr.2020.102803

Ikart, E. (2019). Survey questionnaire survey pretesting method: An evaluation of survey questionnaire via expert reviews technique. *Asian Journal of Social Science Studies, 4(2)*. https://doi.org/10.20849/ajsss.v4i2.565

Jaya, S., Zaharudin, R., Hashim, N. A., Zaid, S. M., Ithnin, M. A., Mapjabil, J., & Nordin, M. N. (2021). Employing design and development research (DDR) approach in designing next generation learning spaces (NGLS) in teachers' pedagogy and technology tools. *Review of International Geographical Education Online, 11*(7), 1237-1246. https://doi.org/10.48047/rigeo.11.07.116

Jhangiani, R.S., Chiang, I. A., Cuttler, C., & Leighton, D.C. (2019). Research methods in psychology (4th ed.). Kwantlen Polytechnic University. https://kpu.pressbooks.pub/psychmethods4e/

Julie, H., Sanjaya, F., & Anggoro, A. Y. (2017). The students' ability in the mathematical literacy for uncertainty problems on the PISA adaptation test. *AIP Conference Proceedings, 1868(1):050026.* https://doi.org/10.1063/1.4995153

Kablan, Z. & Uğur, S. S. (2019). The relationship between routine and non-routine problem solving and learning styles. *Educational Studies, 1-16.* https://doi.org/110.1080/03055698.2019.1701993

Kartikasari, I. A., & Usodo, B. (2022). The Effectiveness open-ended learning and creative problem solving models to teach creative thinking skills. *Pegem Journal of Education and Instruction*, *12*(4), 29-38.. https://doi.org/ 10.47750/pegegog.12.04.04

Kinnear, G., Jones, I., Sangwin, C., Alarfaj, M., Davies, B., Fearn, S., ... & Wong, T. (2024). A collaboratively-derived research agenda for e-assessment in undergraduate mathematics. *International Journal of Research in Undergraduate Mathematics Education*, *10*(1), 201-231. https://doi.org/10.1007/s40753-022-00189-6

Kolovou, A., van den Heuvel-Panhuizen, M., & Bakker, A. (2009). Non-routine problem solving tasks in primary school mathematics textbooks – a needle in a haystack. *Mediterranean Journal for Research in Mathematics Education, 8 (2), 31-68.* https://Kolovou-etal_2009_MJRME_Textbook-analysis%20.pdf

Kragt, D., & Day, D. V. (2020). Predicting leadership competency development and promotion among high-potential executives: The role of leader identity. *Frontiers in psychology, 11, 1816.* https://doi.org/10.3389/fpsyg.2020.01816

Kumari, S., Gautam, H., Nityadarshinin, N., Das, B. K. & Chaudhry. (2021). Online classes versus traditional classes? Comparison during COVID-19. *Journal of Education and Health Promotion, 10, 1-4.* https://doi.org/10.4103/jehp.jehp_317_21

Kundu, A. & Bej, T. (2021). Experiencing e-assessment during COVID-19: an analysis of Indian students' perception. *Higher Education Evaluation and Development, 15(2), 114-134.* https://www.emerald.com/insight/2514-5789. htm

Kurniawan, H., Darmono, P. B., Mursalin, M., Shang, Y., Weinhandl, R., & Sharm, R. (2022). Describe mathematical creative thinking skills and problem-solving strategies by prospective teacher students on non-routine problems. *International Journal of Trends in Mathematics Education Research*, *5*(2), 119-124. https://doi.org/10.33122/ijtmer.v5i2.140

Lee, H. Y., Chen, P. H., Wang, W. S., Huang, Y. M., & Wu, T. T. (2024). Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, *21*(1), 1-28. https://doi.org/10.1186/s41239-024-00447-4

Legarde, M. A. (2022). Students' common errors in solving routine & non-routine problems: a mixed method analysis. *International Journal of Multidisciplinary Research and Analysis, 5(2), 541-547.* https://doi.org/10.47181/ijmra/v5-i2-42

Liebenberg, L., Jamal, A., & Ikeda, J. (2020). Extending youth voices in a participatory thematic analysis approach. International Journal of Qualitative Methods, 19, 1609406920934614. https://doi.org/10.1177/1609406920934614

Makmuri, Aziz, T. A. & Kharis, S. A. A. (2021). Characteristics of problems for developing higher-order thinking skills in mathematics. *Journal of Physics: Conference Series, 1-9.* https://doi.org/ 10.1088/1742-6596/1882/1/ 012074

Manopo & Lisarani, V. (2021). A Comparative analysis of non-routine problems in mathematics textbooks of Indonesia and Singapore. *EDUMATIKA: Jurnal Riset Pendidikan Matematika, 4(1).* https://doi.org/10.32939/ejrpm.v4i1.750

Marriott, P. (2009). Students' evaluation of the use of online summative assessment on an undergraduate financial accounting module. *British Journal of Educational Technology, 40(2), 237-254.* https://doi.org/10.1111/j.1467-8535.2008.00924.x

Mogari, D., & Lupahla, N. (2013). Mapping a group of Northern Namibian grade 12 learners' algebraic non-routine problem solving skills. *African Journal of Research in Mathematics, Science and Technology Education, 17(1/2), 94-105.* https://doi.org/10.1080/10288457.2013.826974

Monrat, N., Phaksunchai, M., & Chonchaiya, R. (2022). Developing students' mathematical critical thinking skills using open-ended questions and activities based on student learning preferences. *Education Research International, 2022, Article ID 330036.* https://doi.org/10.1155/2022/3300363

Muhammad, N., Ozuem, W., Howell, K., & Ranfagni, S. (2023). A Step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *International Journal of Qualitative Methods, 22*, 1–18. https://doi.org/10.1177/16094069231205789

Padzil, M.R., Karim, A.A., & Husnin, H. (2021). Employing DDR to design and develop a flipped classroom and project based learning module to applying design thinking in design and technology. *International Journal of Advanced Computer Science and Applications, 12*(9), 791-798. https://doi.org/10.14569/IJACSA.2021.0120988

Pamungkas, A.P., & Widodo, P. (2019). Implementing online course evaluation rubric on e-learning ecologies: innovative approaches to teaching and learning for the digital age. *Proceedings of the International Conference on Online and Blended Learning 2019 (ICOBL 2019) volume 440, 138-141.* https://doi.org/10.2991/assehr.k.200521.030

Pegg, J. (2010). Promoting the acquisition of higher order skills and understandings in primary and secondary mathematics, In Research Forum. https://research.acer.edu.au/cgi/viewcontent.cgi?article=1081&context=research_conference

Pham, A. T. (2022). University students' attitudes towards the application of quizizz in learning english as a foreign language. *International Journal of Emerging Technologies in Learning, 17(19).* http://doi.org/10.3991/ijet.v17i19.32235

Pitorini, D. E., Suciati, & Harlita (2024). Students' critical thinking skills using an e-module based on problem-based learning combined with socratic dialogue. *Journal of Learning Development, 11*(1), 52-65. https://doi.org/10.56059/jl4d.v11i1.1014

Plomp, T. (2007). Educational design research: an introduction. an introduction to educational design research. *Proceedings of the seminar conducted at the East China Normal University, Shanghai (PR China), November 23-26, 2007, 9-35.* https://ris.utwente.nl/ws/portalfiles/portal/14472302/Introduction_20to_20education_20design_20research.pdf

Rasila, A., Malinen, J., & Tiitu, H. (2015). On automatic assessment and conceptual understanding. Teaching Mathematics and its Applications, 34(3), 149—159. https://doi.org/10.1093/teamat/hrv013

Richey, R. C., & Klein, J. D. (2007). *Design and development research: methods, strategies, and issues.* https://openlibrary.org/books/OL11281240M/Design_and_Development_Research

Rusmini, R., Suyono, & Agustini, R. (2021). Profile of Argumentation ability of undergraduate students in chemistry education based on non-routine problems. *E3S Web of Conferences.* https://doi.org/10.1051/e3sconf/ 202132806007

Sahrir, M. S., Alias, N. A., Ismail, Z., & Osman, N. (2012). Employing design and development research (DDR): approaches in the design and development of online arabic vocabulary learning games prototype. *Turkish Online Journal of Educational Technology, 11*(2), 108-119. Retrieved from https://files.eric.ed.gov/fulltext/EJ989018.pdf

Saro, J. M. et al. (2022). New normal education: strategies, methods, and trends of teaching-learning on students' perspectives and effectiveness. *Psychology and Education: A Multidisciplinary Journal, 5(1), 259-265.* https://doi.org/10.5281/zenodo. 7242770

Shawan, M., Osman, S. & Abu, M. S. (2021). Difficulties in solving non-routine problems: preliminary analysis and results. *ASM Science Journal, 16, 1-14.* https://doi.org/10.32802/asmscj.2021.800

Sikurajapathi, I., Henderson, K., & Gwynllyw, R. (2020). Using e-Assessment to Address Mathematical Misconceptions in Engineering Students. *International Journal of Information and Education Technology, 10(5), 356-361.* https://doi.org/10.18178/ijiet.2020.10.5.1389

Su, C. Y. & Guo, Y. Q. (2021). Factors impacting university students' online learning experiences during the COVID-19 epidemic. *Journal of Computer Assisted Learning, 37, 1578-1590.* https://doi.org/10.1111/jcal.12555

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. American Journal of Evaluation, 27(2), 237–246. https://doi.org/10.1177/ 1098214005283748

Tonbuloglu, B. (2022). State of research on e-assessment in education: a bibliometric analysis. *Turkish Online Journal of Distance Education-TOJDE, 24(4), 276-301.* https://doi.org/10.17718/tojde.1184218

Truong, K. D. & Tran, B-C. N. (2021). Higher order thinking skills in teaching academic writing: suggestions for application. *Theodore Maria School of Arts, Assumption University, 101-124.* http://www.assumptionjournal.au.edu/index.php/newEnglishTeacher/article/view/ 5581/3329

Usta, N. (2020). Evaluation of preservice teacher's skills in solving non-routine mathematical problems through various strategies. *Asian Journal of Education and Training, 6(3), 362-383.* https://doi.org/10.20448/ journal.522.2020.63.362.383

Van der Kleij, F. M., Feskens, R.C.W., & Eggen, T.J.J.M. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: a meta-analysis. *Review of Educational Research, 85(4), 1-37.* https://doi.org/10.3102/0034654314564881

Van Zanten, M. & van den Heuvel-Panhuizen, M. (2018). Opportunity to learn problem solving in Dutch primary school mathematics textbooks. *ZDM.* https://doi.org/10.1007/s11858-018-0973-x

Wang, J. (2011). Re-examining test item issues in the TIMSS Mathematics and Science Assessments. *School Science and Mathematics, 111(7), 334-344.* https://doi.org/10.1111/j.1949-8594.2011.00096.x

Wang, Y., Kai, S. & Baker, R. S. (2020). Early detection of wheel-spinning in assistments. *International conference on artificial intelligence in education, 574-585.* https://doi.org/10.1007/978-3-030-52237-7_46

Wijaya, T., Hidayat, W., & Zhou, Y. (2020). Development of Interactive Learning Video on Linear Program. *Universal Journal of Educational Research, 8(12A), 7530 - 7538.* https://doi.org/ 10.13189/ujer.2020.082537

Xu, W., &Zammit,K.(2020).Applying thematic analysis to education: A hybrid approach to interpreting data in practitioner research. Inter national Journal of Qualitative Methods, 19, 1609406920918810. https://doi.org/10.1177/1609406920918810

Yazgan, Y. (2015). Sixth graders and non-routine problems: Which strategies are decisive for success? *Educational Research and Reviews, 10(13), 107-1816.* https://doi.org/10.5897/ERR2015.2230

Yolida, B., Marpaung, R. R. T., & Lengkana, D. (2023). Pre-service teachers': analysis and constraints in making hots-based multiple choice questions. *Journal of Education for Sustainability and Diversity, 2(1), 59-79.* https://doi.org/10.57142/jesd.v2i1.59

Yu, R. & Cai, X. (2022). Impact of immediacy of feedback on continuous intentions to use online learning from the student perspective. *Frontiers in Psychology, 13(865680).* https://doi.org/10.3389/fpsyg.2022.865680

Zakaria, Z. A., Abdullah, N. & Mahir, N. A. (2022). The perceptiosn of students on online assessment during covid-19 pandemic. *International Journal of Languages and Education, 1(1), 65-73.* https://doi.org/xxx.0000/jcicom.vol1no1.9

Zulkipli & Saidah, M. (2024). How do junior high school students solve problems in non routine mathematics problems. *Journal of Education for Sustainability and Diversity, 3(1), 65-76.* https://doi.org/10.57142/jesd.v3i1.616