International Electronic Journal of Mathematics Education
2024, 19(3), em0787
e-ISSN: 1306-3030
https://www.iejme.com

MODESTUM

Research Article

OPEN ACCESS

Development and effectiveness of digital classroom assessment document for form one mathematics in secondary school

Nurihan Nasir [1*], Mazlini Adnan [1], Murugan Rajoo [1], Anis Oweeda Ismail [2], Riyan Hidayat [3,4]

[1] Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, Perak, MALAYSIA
[2] Pejabat Pendidikan Daerah Kerian, Parit Buntar, Perak, MALAYSIA
[3] Faculty Educational Studies, Universiti Putra Malaysia, Selangor, MALAYSIA
[4] Institut Penyelidikan Matematik, Universiti Putra Malaysia, Seri Kembangan, MALAYSIA
***Corresponding Author:** nurihan@fsmt.upsi.edu.my

Citation: Nasir, N., Adnan, M., Rajoo, M., Ismail, A. O., & Hidayat, R. (2024). Development and effectiveness of digital classroom assessment document for form one mathematics in secondary school. *International Electronic Journal of Mathematics Education, 19*(3), em0787. https://doi.org/10.29333/iejme/14766

ARTICLE INFO

Received: 06 Jan. 2024

Accepted: 10 Jun. 2024

ABSTRACT

Classroom assessment is essential for tracking students' progress and improving teaching and learning in the classroom. However, the lack of clear documentation to guide teachers in assessing student mastery often hinders effective communication between teachers and stakeholders about the students' progress. This study aimed to develop and test the digital classroom assessment document (CA-Do) for form one mathematics to improve classroom assessment practices. Design research and development method, based on ADDIE model, was used to create the digital CA-Do, and 55 mathematics teachers from 55 schools were sampled using stratified sampling. The data collected was then analyzed using SPSS version 27. The digital CA-Do received a satisfactory expert validity score of content validity index (1.00), and found all four tested sub-constructs (i.e., usefulness, ease of use, ease of learning, and satisfaction of use) to be of high level. There were significant differences in teachers' knowledge post-familiarization with the digital CA-Do. The study concludes that the digital CA-Do can aid teachers in recording the intervention and impact of teaching and learning information and consistently assessing pupils' mastery level. This research provides guidance to educators on how to modify their classroom assessment strategies to enhance teaching outcomes and classroom assessment methods, particularly formative assessment.

Keywords: classroom assessment, digital classroom assessment document, intervention, mastery level, usability

INTRODUCTION

Assessment is an important component of education that involves collecting, analyzing (Baidoo-Anu et al., 2023; Balbi et al., 2022; Fauziah et al., 2018), and interpreting data to evaluate student learning outcomes and inform teachers about instructional outcomes (Balbi et al., 2022; Black & Wiliam, 2018). One of the most important functions of assessment is to help teachers measure student knowledge, skills, abilities, and determine student progress by learning objectives (Brandmo et al., 2020). It can also help teachers assess student learning (Mohd Isa et al., 2020), identify student problems, and find the best solutions to improve the quality of teaching (Foster, 2022). Assessment can take various forms, including tests, quizzes, projects, essays, and observations, and also can be used for various purposes, including grading, feedback and accountability (Boström & Palm, 2023; Haj-Yahya & Olsher, 2022; Kamarudin et al., 2021).

Currently, various forms of assessment are used by teachers to assess student learning outcomes, such as formative assessment, authentic assessment, portfolio assessment, peer assessment, and so on. The results of previous studies have shown that most assessments used in schools can measure some of the important skills that pupils desperately need to improve their attainment. This finding can be evidenced by Kultur and Kutlu (2021), which show that formative assessments positively impact pupils' achievement and attitudes toward mathematics when they learn a topic well. Other findings also show that formative assessments increase student engagement, improve learning outcomes (Foster, 2022; Kultur & Kutlu, 2021; Rumanová et al., 2020), increase student motivation and develop a deeper understanding of mathematical concepts, and improve pupils' skills and increase their confidence in their abilities (Foster, 2022). In addition, peer assessment practices have increased student engagement and academic achievement and helped pupils develop critical thinking, communication, and teamwork skills (Sanaeifar & Mirshojaee, 2020). While authentic assessment can improve student learning, it is an alternative to help pupils understand a concept well (Balbi et al., 2022; Fauziah et al., 2018). Other findings related to portfolio assessment have shown that the progress of pupils assessed individually could be well measured and used to effectively improve student learning, enhance the quality of education, and promote positive student success (Brandmo et al., 2020; Lestariani et al., 2018).

Although the existing assessment gives a positive indication of the student, the deficiencies still need to be improved to make it more comprehensive, holistic, and fair in assessing the student's mastery level. This indication can be evidenced by the study of Arumugham (2020), which asserts that the existing assessment in schools is not holistic. This finding is because the existing assessment cannot optimally measure educational outcomes, and an alternative assessment is needed to serve as a supportive assessment that can improve educational outcomes. This statement is also supported by problems in the Malaysian education system that require alternative assessment to solve the problems that arise, especially in mathematics. To overcome this problem, the Malaysian Ministry of Education introduced the level one education transformation in 2011, which aims to improve the quality of education by moving from a traditional examination-based assessment system to an assessment approach. This approach is known as the more formative classroom assessment, which focuses on student development in the classroom. This development shift aims to increase teacher autonomy in the classroom and promote a more student-centered educator approach (Mohd Isa et al., 2020).

According to Zamri and Hamzah (2019), classroom assessment is an assessment approach used in classrooms in Malaysia that aims to evaluate pupils' academic performance based on the learning experience in the classroom. The implementation of classroom assessment also aims to provide meaningful feedback on student's progress and identify areas, where they may need additional support or guidance from teachers. In addition, classroom assessment is important for helping teachers understand their pupils' strengths and weaknesses so that they can reflect on their teaching after the intervention is implemented. In addition, classroom assessment can fairly and consistently measure student progress, identify problem areas, and find the best solutions to improve student achievement. To effectively implement classroom assessment, several requirements need to be met. First, teachers must thoroughly understand the proximity of assessment and ensure that it aligns with their teaching goals. Teachers must also be trained to administer the assessments and accurately interpret the results (Zamri & Hamzah, 2019). Second, the assessment tools used for classroom assessment must be valid and reliable to ensure that measures of student learning are accurate. Finally, a feedback mechanism should allow teachers to provide pupils and parents timely feedback on student progress (Hidayat et l., 2023; Mohd Isa et al., 2020).

It is important to remember that while classroom assessment is one of the good assessment approaches, it should not be the only pillar to assess student performance. Teachers should use different assessment methods to comprehensively understand their pupils' learning progress (Arumugham, 2020; Mohd Isa et al., 2020; Syaifuddin, 2019). In addition, it is important to create a positive and supportive classroom climate that focuses on student growth and not just grades. In this way, pupils can develop an interest in learning and a desire to self-indulge (Guarella et al., 2022). However, implementing classroom assessment has raised several issues, including concerns about the impact on teacher autonomy. Similarly, for instance, mathematics educators at the school level exhibit limited familiarity with contemporary digital resources of the 21st century (Hidayat et al., 2024; Joshi & Rawal, 2021). Some educators believe that the new system places too much emphasis on standardized testing and limits their ability to tailor lessons to pupils' needs. Other concerns include that the move away from traditional exams will likely lead to less academic assertiveness and accountability. In addition, the assessment process of classroom assessment itself has been criticized. Some argue that it does not adequately measure student learning and places a burden on teachers that should not happen (Mohd Isa et al., 2020).

For this reason, researchers have taken the initiative to create a digital classroom assessment document (CA-Do) of form one mathematics to evaluate the assessment process in the classroom. Joshi et al. (2023) highlight the significance of enabling teachers with technology to improve their proficiency and expertise in utilizing digital tools. This empowerment enables teachers to efficiently convert conventional mathematics classrooms into online settings that offer greater advantages for students. The research initiative is also supported by the results of the literature studies, which show that there is no empirical evidence on the construction of CA-Do with holistic characteristics. These characteristics can facilitate teachers to more effectively capture and assess pupils, indirectly improving the quality of instruction and student achievement. The findings of Zamri and Hamzah's (2019) study can also support this research, as research gaps show that teachers do not yet have clear and detailed guidelines for implementing classroom assessment. It is difficult to objectively and consistently assess and verify the results of classroom assessment. The problem of understanding and awareness of classroom assessment among teachers and pupils supports this statement. Therefore, conducting studies on CA-Do of form one mathematics in secondary schools is appropriate.

Research Questions

This study was conducted to answer the following questions:

1. Is there a need to create a CA-Do for form one mathematics?
2. Does CA-Do produced for form one mathematics have satisfactory validity?
3. Is CA-Do of form one mathematics constructed to have satisfactory usability among form one teachers?
4. Is there a significant difference in the knowledge level of form one teachers before and after using CA-Do of form one mathematics?

	Where learner is going	Where learner is right now	How to get there
Teacher	**1** Clarifying learning intentions & criteria for success	**2** Engineering effective class-room discussions & other learning tasks that elicit evidence of student understanding	**3** Providing feedback that moves learners forward
Peer	Understanding & sharing learning intentions & criteria for success	**4** Activating students as instructional resources for one another	
Learner	Understanding learning intentions T criteria for success	**5** Activating students as owners of their own learning	

Figure 1. Five aspects of formative assessment from Black and Wiliam (2009) and Wiliam and Thompson (2008)

LITERATURE REVIEW

Classroom Assessment

Classroom assessment plays a crucial role in gauging students' understanding, informing instructional decisions, and promoting learning (Baidoo-Anu et al., 2023; Balbi et al., 2022). Research supports the idea that well-designed assessments contribute to improved student outcomes (Shepard, 2000). In their comprehensive review, Black and Wiliam (1998) emphasized the formative nature of assessment, highlighting its potential to enhance learning when integrated seamlessly into the teaching and learning process. They argued that effective formative assessment involves ongoing feedback, student involvement, and adjustments to instruction based on assessment results. Besides, Gezer et al. (2021) encourage educators to use formative assessment techniques more regularly in primary school classrooms with low achievement levels. The cycle of data collecting, data analysis, preparing instruction for the future, and evaluating the impact of that instruction by going back to data collection is all part of the formative assessment process.

Furthermore, Brandmo et al. (2020) underscores the significance of fair and valid assessment practices. They discussed the importance of aligning assessments with instructional objectives, ensuring that assessments accurately measure what students are expected to learn. In the realm of technology-assisted assessment, research by Fjørtoft (2020) explored the benefits of incorporating technology in assessment practices. He highlighted the potential of technology to provide immediate feedback, tailor assessments to individual student needs, and enhance engagement. These citations collectively underline the importance of thoughtfully designed, ongoing, and technology-enhanced assessments in the classroom, emphasizing their role in promoting student learning and achievement. Many types of classroom assessments can be conducted in teaching and learning. For example, classroom assessment techniques (Veldhuis & van den Heuvel-Panhuizen, 2019), peer assisted reflection (Reinholz, 2015), multimodal digital classroom assessments (Fjørtoft, 2020), etc.

In the seminal work of Black and Wiliam (1998), the formative nature of assessment takes center stage. They argue that effective formative assessment extends beyond measuring what students know; it actively shapes the learning process. Through ongoing feedback, student involvement, and adjustments to instruction based on assessment results, formative assessment becomes a powerful tool for enhancing learning (Black & Wiliam, 2009, 2018; Wiliam & Thompson, 2008) (**Figure 1**). In previous study, by Zulliger et al. (2022), emphasizing that formative assessment is about guiding teachers and students on the next steps in the learning journey. However, Hattie and Timperley's (2007) meta-analysis on feedback underscores its transformative impact on student achievement. Their findings highlight that timely and specific feedback is a powerful tool for enhancing learning outcomes. The importance of feedback that focuses on the task, the process, and the self-regulation of the learner, as advocated by Black et al. (2004), adds another layer of depth to the discussion. Constructive feedback emerges as a critical element of the assessment process, providing valuable insights for both students and educators. Students are responsible for working independently on a difficult homework problem and provide written feedback and then conference on their solutions (Rumanová et al., 2020).

In the digital age, Fjørtoft (2020) research explores the benefits of incorporating technology in assessment practices. Technology becomes a catalyst for immediate feedback, tailoring assessments to individual student needs, and enhancing overall engagement. This aligns with the evolving landscape of education, where technology serves as a valuable ally in creating more dynamic and responsive assessment practices. In addition, five formative assessment techniques were implemented using eight analogue and digital resources that were identified by Staberg et al. (2022). The methods most frequently employed had to do with "engineering effective classroom discussions" that produced proof of student comprehension and "activating students" to become self-directed learners and peer educators. The main justifications offered by the educators for utilizing the chosen materials had to do with their efficacy, usefulness, and relevance. Discussions of the pedagogical implications and descriptions of teacher interactions with the chosen resources are provided.

According to previous study findings, there are important assessment-related concerns in the field of educational practices that demand attention from all parties involved in ensuring the quality of the teaching and learning process (Akayuure, 2021). Studies have shown that making errors and mistakes throughout the learning process can aid in memory retention and facilitate the application of effective strategies for addressing accomplishment obstacles in the future. A summary of three research streams–thoughts, the nervous system, and self-control that support the need for error is shown in classroom assessment that enhances student learning and motivation (McMillan & Moore, 2020). Daily classroom assessments have latent power since they can help students improve their performance and develop a deeper comprehension of the material (Fjørtoft, 2020; Zhao et al., 2019). In summary, there has been a recent change in the way that teachers and students jointly study assessment perspectives, emphasizing how these perspectives affect classroom culture and learning, highlighting important issues in education, supporting the benefits of making mistakes in the learning process, and highlighting the potential of regular classroom assessments to improve student performance and deepen learning.

METHODOLOGY

Research Design

The study used a design research and development (DRD) supported by a quantitative approach (Hidayat et al., 2020, 2021; Mohd Tajudin et al., 2022). This study used DRD to build CA-Do of form one mathematics. Meanwhile, the quantitative approach by the survey was used to collect research data through questionnaire instruments and knowledge tests that were modified according to the suitability of the purpose of the study. According to Siraj et al. (2021), the type of DRD is a systematic method involving three main phases: the need analysis, design and development, and evaluation. The following is a detailed explanation of each phase used in this study.

Phase 1: Need analysis

This phase was the first to identify the need to build CA-Do of form one mathematics. In this phase, researchers obtained information related to issues or problems in Malaysia's education system, especially in the classroom assessment, by examining the existing classroom assessment with the requirements of the classroom assessment that should be created. The researchers obtained this information based on the literature study's findings and feedback from several mathematical education and assessment experts in district education office (DEO).

Phase 2: Design & development

In this phase, the researchers implemented ADDIE model as a guide in preparing CA-Do of form one mathematics. Generally, this phase was divided into two stages, namely, the design and development stages. Both levels had their functions required by researchers in developing the digital CA-Do mathematics form one. In this study, the design stage served to determine the unique characteristics required by teachers to make it easier for them to record the progress of student's achievement based on the mastery standard and assess the level of mastery (LM) of the pupils. Meanwhile, the development stage was a phase that considered the findings in phase one to develop CA-Do of form one mathematics. The following information are the five stages used in ADDIE model, as follows:

A: Analysis: Level analysis in ADDIE model used by researchers was to analyze the standard curriculum and assessment document (SCAD) of form one mathematics (Ministry of Education Malaysia, 2015, 2019) involving five areas of learning. The fields were numbers and operations, association and algebra, syllabus and geometry, discrete mathematics and statistics and probability. In addition, the researchers also analyzed the second edition of the classroom assessment implementation handbook (Ministry of Education Malaysia, 2015, 2019) and the student development form obtained from selected secondary schools.

D: Design: At this level, CA-Do of form one mathematics was built in line with the secondary school curriculum with several elements compatible with the characteristics required in the assessment in the classroom. The reorganization of this information was a key feature in the design of CA-Do of form one mathematics. The generation of elements resulted from the findings of the analysis in the first stage, which involved the process of restructuring the planning of the content standards (CS) and learning standards (LS) based on performance standards (PS). Next, the researchers emphasized the assessment element by providing the overall assessment summary according to the title, combining the learning standard form and the mastery level form. This summary should make it easier for teachers to track student achievement progress based on CS and to evaluate pupils' LM more effectively.

D: Development: The development stage involved a combination of findings from the previous stage of constructing CA-Do of form one mathematics (**Figure 2**). Based on the design stage in the process of reorganizing CS and LS planning based on PS, the researchers restructured the basic content of the simple to difficult lessons according to the Bloom's taxonomy skill level. They linked with the contents of the form one mathematics textbook. After the compilation of the learning contents, the focus of this design continued with the development of the teacher's teaching by looking at the number of pupils that CS achieves after the teacher's teaching and after the teacher has intervened. In addition, there were blank spaces or intervention notes provided for improving teaching by form one mathematics teachers. The final part of the chapter contained a summary of the entire assessment according to the chapter and the overall mastery levels of form one mathematics. According to the title, the assessment results were formulated in the list of the overall LM of form one mathematics in the final part of this CA-Do.

I: Implementation: The researchers conducted a pilot study in three Muallim District Perak State secondary schools. This pilot study involved 25 mathematics teachers of form one to improve CA-Do of form one mathematics. This pilot study identified strengths and weaknesses in the form one mathematical digital CA-Do. Based on the findings of this pilot study, researchers did make improvements to address the weaknesses when implementing this assessment document in the classroom. The exposure

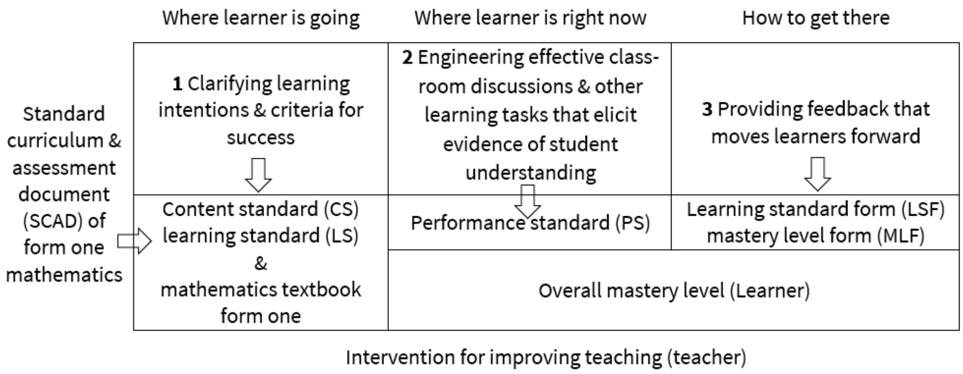

Figure 2. Three aspects of formative assessment (1, 2, & 3) adapted from Black and Wiliam (2009) & Wiliam and Thompson (2008) in development framework of digital CA-Do mathematics form one

Table 1. District education office by zone

Zone	Area of district education office	Number of DEO	Selected secondary schools
North	Kerian DEO, Larut Matang & Selama DEO, Kuala Kangsar DEO, & Hulu Perak DEO	4	20
Middle	Kinta Utara DEO, Kinta Selatan DEO, & Perak Tengah DEO	3	17
South	Muallim DEO, Batang Padang DEO, Bagan Datuk DEO, Hilir Perak DEO, & Manjung DEO	5	19
Total		12	55

of form one mathematical digital CA-Do among mathematics teachers who teach form one can improve the quality of this classroom assessment.

E: Evaluation: The assessment stage was the final stage in ADDIE model, used to test the validity of CA-Do of form one mathematics build. The digital CA-Do received the validity of seven experts with experience in mathematics education and assessment. Thus, the researchers considered the feedback from the appointed experts to improve the construction quality of the digital CA-Do of form one mathematics. Overall, the results showed that all experts rated the validity of the researchers' constructed assessment documents as satisfactory.

Phase 3: Valuation

This phase was the third phase in DRD, which involved two evaluations, namely the usability and effectiveness of the built form one mathematical digital CA-Do. The usability assessment used in this study aimed to determine the level of usability of the form one mathematical digital CA-Do built at low, medium, or high levels by involving the four main sub-constructs tested: usefulness, ease of use, ease of learning, and satisfaction. Meanwhile, the evaluation related to the effectiveness of the digital CA-Do mathematics form one was created to determine if there were significant differences in the knowledge of the mathematics teacher form one before and after instruction in the digital CA-Do of form one mathematics. After discussion with the research partners, they agreed that the assessment of the effectiveness of the form one mathematical digital CA-Do only included one group of form one mathematics teachers. This is because the COVID-19 pandemic in this country still has a trend of unstable infection rates that could indirectly affect the intervention in this study.

Population & Sampling

The study population consisted of form one mathematics teachers in Perak, Malaysia. The 12 districts in the state of Perak comprise the northern, middle, and southern zones. Each district has its own DEO, responsible for managing primary and secondary schools. **Table 1** shows the number of DEOs in the northern, middle, and southern zones. The sample for the study consisted of the number of teachers randomly selected from DEOs. The selected DEO is based on each Northern Zone, Middle Zone, and Southern Zone in the State of Perak, Malaysia.

The study was conducted in three DEOs, namely Kuala Kangsar DEO with 20 secondary schools, Perak Tengah DEO with 17 secondary schools, and Manjung DEO with 19 secondary schools. Each school is represented by a form one mathematics teacher who is permitted each DEO. The sample selection included 55 form one mathematics teachers from 55 secondary schools assigned to each zone in the state of Perak, Malaysia. **Table 2** summarizes the profile information of the form one mathematics teachers selected as the sample for the study.

Instruments

The study used three main instruments, namely CA-Do validity questionnaire (CA-DoV), CA-Do usability questionnaire (CA-DoU), and the teacher knowledge test (TKTe). These three instruments were used as data collection tools by researchers based on research to build and study the level of usability and effectiveness of CA-Do of form one mathematics. In this study, the digital CA-DoV item was built based on the guidebook on the Implementation of design and development studies and the design of development studies written by Mohd Tajudin et al. (2022), which has been modified according to the needs and suitability of the study. The construction of facial validity items was intended to measure the extent to which the measurement tool can measure the actual value it wants to measure (Nunnally & Bernstein, 1994). In addition, facial validity also refers to the suggestions for

Table 2. Respondents profile information-1

Item	Category	Frequency (n=55)	Percentage (%)
Gender	Male	13	23.64
	Female	42	76.36
Age	20-29	6	10.91
	30-39	14	25.45
	40-49	24	43.64
	50-59	11	20.00
DEO	Kuala Kangsar DEO	18	32.73
	Manjung DEO	20	36.36
	Perak Tengah DEO	17	30.91
Duration of service	1-5	6	10.91
	6-10	5	9.10
	11-15	14	25.45
	16-20	15	27.27
	21-30	14	25.45
	31+	1	1.82
Highest approval	Diploma/specialization	0	0.00
	Degree	51	92.73
	MSc	4	7.27
	PhD	0	0.00
Study options	Mathematics	35	63.64
	Non-mathematical	20	36.36
School location	City	21	38.18
	Rural	34	61.82

Table 3. Respondents profile information-2

Construct	Item number	Item
Face validity	7	1 to 7
Content validity	12	8 to 19
Number of items	19	

Table 4. Item of each assessed sub-construct

Sub-constructs	Item number	Item
Usefulness	8	1 to 8
Ease of use	10	9 to 18
Ease of learning	7	19 to 25
Satisfaction of use	7	26 to 32
Number of items	32	

improvement given by experts based on the appearance of the instrument considering aspects of consistency of style and format, clarity to read, and clarity of language (Hidayat et la., 2018; Taherdoost, 2016).

Meanwhile, constructing content validity items was intended to determine whether the items represented by a construct can measure the characteristics the construct should measure. In other words, a measurement should include the entire domain of the construct studied (Wynd et al., 2003). The instrument consists of two main sections: part A and part B. Part A details the respondent's profile information, such as gender, age, length of service, teaching experience, highest graduation, study options, and school location. While part B submitted 19 items of assessment of the constructed document, namely digital CA-Do mathematics form one using a four-point Likert scale, where this instrument requires the respondent to mark only one answer option, which is '1-irrelevant', '2-less relevant', '3-relevant', and '4-very relevant'. **Table 3** refers to items used in CA-DoV instruments:

Next, the item usability questionnaire (CA-DoU) was adopted from Lund's (2001) study and modified according to the needs and suitability of this study. The constructed item consists of two main components to collect the study data. The first component (part A) contains respondent profile information such as gender, age, length of service, teaching experience, highest graduation, study opportunities, and school location, while the second component (part B) contains 32 items with four sub-constructs, namely usefulness, ease of use, learning facilities and satisfaction with use, which are used to determine the level of ease of use of the form one mathematics CA-Do among teachers. For each item, a four-point Likert scale was used on which the expert marks only one valid response option, i.e., '1-irrelevant', '2-less relevant', '3-relevant', and '4-very relevant'. **Table 4** contains information on creating the usability questionnaire items.

The third instrument is TKTe. TKTe was used to determine if there were significant differences in teachers' knowledge before and after exposure to the form one mathematics CA-Do. This instrument consists of subjective questions that require the respondent to answer all questions. The test items created are based on the test specification table (TeST) created beforehand. According to Mohd Tajudin et al. (2022), TeST is the best way to improve the validity of a test. In addition, experts provided feedback or suggestions for improvement for each item in TKTe.

Table 5. Classroom assessment requirement analysis findings

Item	Percent score (%)
1. The implementation of classroom assessment needs to be strengthened.	
Agree	100
Disagree	0.00
2. Teachers have been very burdened since the classroom assessment was implemented.	
Agree	75.00
Disagree	25.00
3. Teachers do not have specific assessment documents to record the level of mastery of pupils in classroom assessment.	
Agree	62.00
Disagree	38.00
4. Teachers do not have appropriate assessment documents to record intervention and impact of teaching and learning.	
Agree	68.00
Disagree	32.00
5. A specific assessment document with simple features maybe developed to make it easier for teachers to implement classroom assessment.	
Agree	93.00
Disagree	7.00
6. If given the opportunity, which is your choice?	
60.00% classroom assessment and 40.00% examination	16.00
40.00% examination and 60.00% classroom assessment	84.00

Data Collection Procedure

Before conducting the study, the researchers obtained a letter of authorization from the department of educational policy planning and research (EPRD) and an approval letter from Universiti Pendidikan Sultan Idris (UPSI) for the ethical evaluation of human subjects research. In addition, the researchers sought permission from the Perak State Education Department (SED) and DEO to conduct the study in the selected schools. After obtaining permission from EPRD and Perak SED, this study was conducted with the first needs assessment phase. The researchers provided 25 secondary mathematics teachers in Kuala Kangsar and Manjong districts with a needs assessment questionnaire. These secondary mathematics teachers are used as respondents to get an initial picture of the need for developing CA-Do of form one mathematics. Then, the data are analyzed using descriptive statistics to identify the elements needed in the digital CA-Do mathematics form one design. Next, the researchers examined the level of usability using descriptive statistics and analyzed the effectiveness of the data of CA-Do using the paired sample t-test.

Pilot Study

Before the actual study, a pilot study was conducted to determine the research rationale, feasibility, and suitability of the study and to increase the validity of the research instrument (Chua, 2014). This statement is supported by Konting (2000), who believes that pilot studies must be conducted before the actual study because the results can determine the validity and reliability of the instrument used. In addition, this process aims to identify potential problems in the study and assess the validity and reliability of an instrument (Cohen et al., 2018). The pilot study was conducted in the southern zone, selected through a simple random sampling procedure by zones in the state of Perak, Malaysia. The selected sample has the same characteristics as the actual sample of the study. A total of 25 secondary mathematics teachers were directly involved in this pilot study as the sample. This number is sufficient for a pilot study (Ary et al., 2019). The researchers conducted this pilot study themselves. The data obtained from this pilot study were analyzed using statistical package for social sciences version 27 software to test the instrument's internal consistency.

The pilot study results for CA-Do usability questionnaire showed that the validity of the study instrument was at the expert agreement level of 0.96, proving that the obtained content validation index (CVI) value shows that the instrument's validity in this study is good. The reliability analysis of CA-Do usability questionnaire showed that the total Cronbach's alpha coefficient value was 0.87 for the four tested sub-constructs, i.e., usefulness, ease of use, ease of learning, and satisfaction with use, with Cronbach's alpha coefficient values of 0.91, 0.88, 0.77, and 0.92, respectively. This clearly shows that the items included in this instrument have excellent reliability. As for the instrument TKTe, the value of the Cronbach's alpha coefficient is also acceptable; that is, the value of the recorded coefficient is 0.74. The analysis results show that the instrument's validity and reliability can be adopted for the actual research since it has a satisfactory validity and reliability value.

FINDINGS

Need Analysis for Classroom Assessment

This study used descriptive statistics to analyze needs analysis data, which was the study's first phase. In this phase, six items were used by the researchers to determine whether or not CA-Do of form one mathematics should be developed. This can be determined by the percentage score for each item created. **Table 5** shows the percentage score for each item in the need analysis questionnaire form.

Table 5 shows that all teachers agree 100% that implementing classroom assessment should be systematically strengthened in schools. This is because the systematic implementation of classroom assessment can increase the validity of the assessment data obtained. However, 75% of teachers think implementing classroom assessment in schools is only a burden for them, while

Table 6. Assessment of I-CVI & S-CVI values for face validity of digital CA-Do mathematics form one-1

Item	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	Expert 7	I-CVI
1	X	X	X	X	X	X	X	1.00
2	X	X	X	X	X	X	X	1.00
3	X	X	X	X	X	X	X	1.00
4	X	X	X	X	X	X	X	1.00
5	X	X	X	X	X	X	X	1.00
6	X	X	X	X	X	X	X	1.00
7	X	X	X	X	X	X	X	1.00
S-CVI								1.00

Table 7. Assessment of I-CVI & S-CVI values for face validity of digital CA-Do mathematics form one-2

Item	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	Expert 7	I-CVI
1	X	X	X	X	X	X	X	1.00
2	X	X	X	X	X	X	X	1.00
3	X	X	X	X	X	X	X	1.00
4	X	X	X	X	X	X	X	1.00
5	X	X	X	X	X	X	X	1.00
6	X	X	X	X	X	X	X	1.00
7	X	X	X	X	X	-	X	0.86
8	X	X	X	X	X	-	X	0.86
9	X	X	X	X	X	-	X	0.86
10	X	X	X	X	X	-	X	0.86
11	X	X	X	X	X	X	X	1.00
12	X	X	X	X	X	X	X	1.00
S-CVI								0.95

Table 8. Minimum stage measurement

Phase	Range
Low	1.00-2.40
Medium	2.41-3.80
High	3.81-5.00

25% think implementing classroom assessment in schools is not a burden. This study also found that 62% of teachers think that teachers do not have the necessary assessment documents to record pupils' LM in classroom assessment. This statement is also supported by the results of the fourth item, which showed that 68% of the teachers reported not having the appropriate assessment documents to capture the student's LM after the teacher intervened. The teacher could not assess the impact of the intervention implemented, whether or not it was effective in teaching and learning (T&L). However, 84% of the teachers agreed to conduct a 40% classroom assessment and administer only 60% of the exams, compared to 16% agreeing to conduct a 60% classroom assessment and holding only 40% of the examinations. The highest percentage results of the needs analysis showed that 93% of teachers need a specific assessment document and simple characteristics that need to be developed to help teachers conduct class assessments. Therefore, the researchers developed an effective CA-Do as evidence of T&L implementation that directly correlates with determining student mastery levels (LM) in the classroom.

Digital Classroom Assessment Document Validity Analysis

In addition to expert instrument validation, validation was conducted using CA-Do of form one mathematics. The experts were appointed based on their specialization in education and mathematics. A total of seven experts were involved in evaluating the digital CA-Do of form one mathematics using CVI method.

Based on **Table 6** and **Table 7**, it is found that all I-CVI and S-CVI values are accepted. Hence, the results of the experts in determining the validity of the face and the validity of the content of CA-Do of form one in CA-DoV are satisfactory. Apart from that, the results of validity by experts have highlighted several suggestions to improve the digital CA-Do of form one mathematics. The researchers considered these suggestions for improvement before conducting the actual study.

Digital Classroom Assessment Document Usability Analysis

Descriptive statistics were used to analyze the usability assessment data of CA-Do of form one, the third phase of this study, and can be obtained by measuring the mean score for each sub-construct created. In this study, the researchers used four sub-constructs of usability, namely usefulness, ease of use, ease of learning, and satisfaction with use, to determine whether the usability of CA-Do is satisfactory among form one teachers in the secondary schools. The feedback obtained from the respondents was based on the following Likert scale: 1-highly disagree, 2-disagree, 3-less agree, 4-agree, and 5-strongly agree.

Table 8 was adapted from Landell (1997) and served as a reference for researchers in determining the level of applicability of CA-Do of form one for each item used to identify whether the usability sub-constructs used in this study are at a low, medium, or high level. **Table 9** shows average usability construct mean score.

Table 9. Average usability construct mean score

Construct	Sub-constructs	Average mean score	Interpretation
Usability	Usefulness	4.21	Good
	Ease of use	4.16	Good
	Ease of learning	4.08	Good
	Satisfaction of use	4.18	Good

Table 10. Findings of sub-constructs of usefulness

No	Item	Mean score	Interpretation
1	Digital CA-Do helped me become more effective in carrying out the assessment.	4.22	Good
2	Digital CA-Do helped me become more productive to produce assessments.	4.16	Good
3	Digital CA-Do helped me implement a comprehensive assessment according to the title.	4.22	Good
4	Digital CA-Do meets my requirements as a teacher to record classroom assessments.	4.27	Good
5	The use of digital CA-Do saves me time to perform the assessment.	4.13	Good
6	Digital CA-Do is very useful for the purpose of improving teaching and learning.	4.25	Good
7	Digital CA-Do gave me the ability to plan teaching and learning interventions.	4.22	Good
8	Digital CA-Do allows me to express the impact on the interventions that have been carried out.	4.18	Good

Table 11. Findings of ease of use sub-construct

No	Item	Mean score	Interpretation
1	Digital CA-Do is easy to use.	4.25	Good
2	Digital CA-Do is concise to use.	4.20	Good
3	Digital CA-Do is user-friendly.	4.13	Good
4	Digital CA-Do helps me design to achieve learning objectives.	4.16	Good
5	Digital CA-Do is flexible in implementing classroom assessments.	4.07	Good
6	The digital CA-Do guidelines are easy to follow.	4.15	Good
7	Digital CA-Do can be used as a reflection document for carrying out follow-up actions (interventions).	4.24	Good
8	Digital CA-Do has a consistent format and structure to enhance teaching and learning.	4.13	Good
9	By using digital CA-Do, I can implement the assessment more effectively.	4.13	Good
10	I can use digital CA-Do when making professional judgments to determine the LM of the pupil.	4.18	Good

Table 12. Findings of sub-constructs of learning facilities

No	Item	Mean score	Interpretation
1	I learned to use digital CA-Do quickly.	4.05	Good
2	I easily studied CS section, LS, & the number of pupils reaching LS included in digital CA-Do.	4.05	Good
3	I easily use the intervention and impact record space in digital CA-Do.	4.13	Good
4	I learned to use the information in LSF to set the pupil's Mastery Level more easily.	4.05	Good
5	The use of LSF and MLF in digital CA-Do is easy to learn.	4.07	Good
6	Use of overall summary of assessment according to the title in digital CA-Do is also easy to understand.	4.15	Good

The analysis of the items for the digital CA-Do usefulness sub-construct, as in **Table 10**, shows that the mean score for each item is at a high level (4.13≤M≤4.27). The fourth item of this sub-construct has the highest mean (mean [M]=4.27) compared to the other seven items. This item shows that teachers need a specific classroom assessment to record the student's mastery level. This is confirmed by the average mean score for the usefulness sub-construct, which is also at the highest level (see **Table 10**), i.e., M=4.21 compared to the other three sub-constructs. This result shows that the usability of the form one mathematical digital CA-Do for the usefulness sub-construct achieves satisfactory usability among teachers in Perak, Malaysia.

Digital CA-Do Usability Analysis for Ease of Use Sub-Constructs

Table 11 shows that the analysis of the items for the sub-construct digital CA-Do ease of use sub-construct is also at a high level, i.e., item one shows the highest mean (M=4.25), and the fifth item shows the lowest mean (M=4.07). Based on the results of these two items, it is clear that digital CA-Do is constructed to be easy to use to record pupils' mastery levels. However, it needs some improvement in flexibility to make it easier for teachers to capture (4.07≤M≤4.25) the information in the classroom assessment. The overall average mean for this sub-construct is M=4.16, as shown in **Table 11**. This mean average indicates that the sub-construct of the ease of digital CA-Do mathematics form one has achieved satisfactory usability among the mathematics teachers of form one in the secondary school.

Digital CA-Do Usability Analysis for Learning Facility Sub-Constructs

Table 12 shows that the analysis of the items for the substructures of learning facilities with classroom assessment in schools is also at a high level of (4.05≤M≤4.15). The six items formed for this sub-construct show that all items have a mean value that reaches a high range. One of the most significant items is the sixth item, which reaches the highest minimum value (M=4.25) compared to the other items. This item explains that with the presence of digital CA-Do, which has special features such as the overall formulation of the assessment according to the title. This is very helpful for teachers to understand pupils' LM in the classroom assessment more easily. However, three items get the same minimum score of M=4.05, which can be interpreted as

Table 13. Findings of usage satisfaction construct finding

No	Item	Mean score	Interpretation
1	I am satisfied with digital CA-Do as a means of improvement in teaching and learning.	4.23	Good
2	I would recommend digital CA-Do to my other friends.	4.18	Good
3	The use of digital CA-Do benefits me.	4.18	Good
4	I found digital CA-Do can help me implement classroom assessment better.	4.24	Good
5	Digital CA-Do can be one of the pieces of evidence of classroom assessment that has been carried out.	4.29	Good
6	All teachers need to have digital CA-Do.	4.27	Good
7	Digital CA-Do is very useful to teachers.	4.33	Good

Table 14. Normality test

Test	Skewness			Kurtosis		
	Statistics	Standard error	Ratio	Statistics	Standard error	Ratio
Pre-test	0.33	0.32	1.03	-0.44	0.63	-0.70
Post-test	-0.51	0.32	-1.59	-0.62	0.63	-0.98

Table 15. Paired samples statistics–Teacher knowledge

Test		Mean	n	Standard deviation
Pair 1	Before	2.65	55	1.57
	After	6.89	55	1.58

relatively weak items. Nevertheless, it still achieves an overall high mean range. However, the specification of this item has some effect on the learning facilities sub-construct, which has the lowest average minimum score in the usability construct compared to the other sub-constructs. This is confirmed by the average mean value of M=4.08 in **Table 12**. However, this sub-construct still achieves satisfactory usability among form one mathematics teachers in the secondary schools.

Digital CA-Do Usability Analysis for Usage Satisfaction Sub-Constructs

Using **Table 13**, the results show that the analysis of the items for the satisfaction sub-construct when using the form one mathematical digital CA-Do build is also at a high level of ($4.18 \leq M \leq 4.33$). The seventh item is the most significant item of this sub-construct to obtain the highest mean (M=4.33) compared to the other items. This item indicates that the teacher agrees that the construct of digital CA-Do can be used as evidence for the teacher during the follow-up (intervention) so that the teacher can reflect on the intervention implemented in the classroom and whether it is effective or ineffective. The second and third items achieved the same mean score of M=4.18, which means that this item is ranked the lowest compared to the other items. Although the minimum scores for these two items are the lowest, researchers can indirectly use these results to improve teachers' specific characteristics for satisfaction with digital CA-Do to reach a maximum level among form one mathematics teachers. The mean average score for the digital CA-Do usage satisfaction sub-construct scored the third highest mean score compared to the other usability sub-constructs. The average mean score is M=4.18 (see **Table 13**), which shows that this sub-construct is high. The results explain that digital CA-Do mathematics form one, specifically designed for the user satisfaction sub-construct, achieves satisfactory usability among form one mathematics teachers in the secondary schools.

Analysis of the Effectiveness of Digital Classroom Assessment Document

Inferential statistics were used to analyze data from the efficacy assessment phase, the third phase of this study, to determine if there are significant differences in teachers' knowledge before and after using the digital CA-Do by form one mathematics. To answer the question about the studies presented, researchers used the paired samples t-test because the study sample included only one group. Before running the paired samples t-test, the researchers reviewed the assumptions that needed to be met before running the test to ensure that the variables used were normally distributed. **Table 14** shows the results of the normality test for the data.

Analysis of Basic Assumptions for Pre- & Post-Test Scores

This study used skewness and kurtosis statistical tests to check the assumptions of the normality test for the data. According to Cohen et al. (2018), two methods can be used to determine whether the data used in a study are normally distributed. First, when looking at the skewness and kurtosis values, the data are normally distributed because the value obtained is close to zero. Second, the data can be considered normally distributed if the ratio of the two statistical tests is between -2 and 2. **Table 14** shows that the ratio values of the statistical test of skewness and the statistical test of kurtosis are between -2 and 2. This explains that the data on teachers' knowledge before and after participating in the form one mathematics CA-Do are normally scattered. Thus, the assumption of the paired samples t-test is satisfied and can be used to answer the question of the fourth study.

Analysis of Paired Sample t-Test Findings

Table 15 shows the means and standard deviations for teachers' knowledge before and after engaging with CA-Do by form one mathematics. Based on the mean and standard deviation, the results show that teachers' knowledge level improved significantly after engaging with digital CA-Do (M=6.89, standard deviation [SD]=1.58) and before engaging with digital CA-Do (M=2.65, SD=1.57, t[54], $p<0.001$). The data also explained that form one mathematics teachers in the secondary school agreed with the researchers' efforts to create a CA-Do with the appropriate characteristics to help teachers capture teaching and learning

Table 16. Paired samples test

	Paired differences			t	df	Sig. (2-tailed)
	Mean	Standard deviation	Standard error			
Pair 1: Before-after	-4.24	1.81	0.24	-17.41	54	<0.001

invention and assess pupils' mastery level more systematically. This is because compared to the mean before and after exposure with the digital CA-Do construct, it clearly shows a significant difference in teacher knowledge mean of 4.24.

Based on **Table 16**, the value of sig. (2-tailed) was used to determine if there were significant differences in teachers' knowledge before and after engaging with the constructed digital CA-Do mathematics form one. The results show that the value of sig. (2-tailed) is less than 0.001 ($p<0.05$). This result indicates that the null hypothesis was successfully rejected. Thus, there is a significant difference in the mean of teachers' knowledge before and after using CA-Do of form one mathematics. In conclusion, these results show that the digital CA-Do is effective among form one mathematics teachers in the secondary school. It is suitable for implementation in secondary schools to help teachers adopt teaching and learning intervention and assess pupils' mastery learning more systematically and comprehensively.

DISCUSSION

The majority of studies conducted in classroom assessment involving formative assessment (Baidoo-Anu et al., 2023; Balbi et al., 2022; Black & Wiliam, 2018; Foster, 2022; Kultur & Kutlu, 2021) authentic assessment (Fauziah et al., 2018), portfolio assessment (Lestariani et al., 2018), and self-assessment (Brandmo et al., 2020) have focused on the role of assessment for students and the goals of assessment in the field of mathematics education. Based on this argument, there are too few studies that develop a form of assessment to assist teachers in assessing learning and determining students' mastery level comprehensively according to the taught mathematical skills on CS and LS in the context of education in Malaysia. This refers to the practice of teachers recording the intervention and impact of teaching and learning information, and consistently assessing pupils' LM. This pertains to the pedagogical practice, wherein educators document the implementation and efficacy of instructional strategies, while continuously evaluating students' level of proficiency. Hence, the purpose of this study was to develop a digital CA-Do mathematics form one that aligns with the assessment requirements of instructors. This study specifically aims to address four specific research questions. In general, the findings of this study offer a favourable indication of the education system in Malaysia through the development of a comprehensive CA-Do.

This document can also be used as evidence for the classroom assessment process to more accurately determine pupils' mastery level based on CS, LS, and PS in SCAD of form one mathematics (Ministry of Education Malaysia, 2015, 2019). In addition, teachers themselves can assess classroom teaching and learning because the teacher does not have a single assessment document that specifically captures or fairly assesses student mastery levels based on the assessment evidence. In this discussion, the researchers focused only on the third and fourth study questions because these two questions indirectly explain the results of the first and second study.

The results of the third study showed that the level of usability of CA-Do of form one mathematics among form one mathematics teachers in the secondary school was satisfactory. The results showed that the four sub-constructs measured in this study, namely usefulness, ease of use, ease of learning, and satisfaction with use, had a mean score above four, which is a high level. Other research found that the overall mean score of usability of classroom assessments was high (Hashim, 2021). The results also explained that the teacher believed that the availability of this CA-Do could further strengthen the assessment process by recording and assessing pupils' mastery levels in the classroom.

The results of the fourth research question revealed a positive indication among the mathematics teachers towards the digital CA-Do mathematics form one. The results showed significant differences in teachers' knowledge after they use the digital CA-Do mathematics form one. Thinwiangthong et al. (2020) showed that teachers' knowledge improved after they conducted formative assessments of teaching and learning. This result also clearly shows that the development of digital CA-Do mathematics form one is needed in all secondary schools in state of Perak, Malaysia. With the existence of digital CA-Do, teachers can conduct better assessments to more effectively and comprehensively evaluate the knowledge level of their pupils according to the topics taught.

CONCLUSIONS

In conclusion, this study answers all the research questions asked in the studies. This is evident from the results of the first research question, which showed that first was form one mathematics teachers in the secondary schools agree with the implementation of CA-Do of form one mathematics construct as an assessment document that can be used as a guide to record and evaluate pupils' mastery level. Besides, it also supports classroom assessment so that teachers can more accurately interpret pupils' expected results from time to time. The results for the second research question showed that the digital CA-Do of form one mathematics build has satisfactory validity among form one mathematics teachers. This indicates that CA-Do was created according to the experts' requirements and is suitable for use in secondary schools. This finding is further corroborated by the outcomes of the investigation into the third research question. The results indicated that the average percentage of the digital CA-Do usability value was deemed satisfactory, exhibiting a high level across all assessed sub-constructs.

Finally, the results for the fourth research question showed a significant difference in the knowledge of form one mathematics teachers in the secondary schools after they were familiarized with the digital CA-Do of form one mathematics. These results suggested that teachers in secondary schools need a structured classroom assessment to implement assessments more systematically and effectively. In addition, the data obtained through this CA-Do can serve as measures for teachers to revisit whether or not the interventions implemented in the classroom are effective. Therefore, it is hoped that stakeholders such as the Ministry of Education Malaysia, SED, and DEO, as well as schools for the survival of education in Malaysia, can work together to further improve the classroom assessment system in classrooms in the country.

Author contributions: NN: conceptualization, formal analysis, funding acquisition, investigation, methodology, software, supervision, writing - original draft, writing review & editing; **MA:** conceptualization, funding acquisition, validation, visualization; **MR:** data curation, funding acquisition; **AOI:** data curation, funding acquisition, investigation, project administration; **RH:** software, supervision, validation, writing - original draft, writing - review & editing. All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: This study was conducted with the support from the Research Management Innovation Center, Universiti Pendidikan Sultan Idris under the Incentive Grant University Research or Geran Galakan Penyelidikan Universiti scheme (code: 2021-0053-107-01).

Ethical statement: The authors stated that they obtained a letter of authorization from the Department of Educational Policy Planning and Research and an approval letter from Universiti Pendidikan Sultan Idris for the ethical evaluation of human subjects research on 28 April 2022 (Approval code: 2022-0542-01). Written informed consents were obtained from the participants.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Akayuure, P. (2021). Classroom assessment literacy levels of mathematics teachers in Ghanaian senior high schools. *Contemporary Mathematics and Science Education, 2*(2), Article ep21013. https://doi.org/10.30935/conmaths/11286

Arumugham, K. S. (2020). Kurikulum, pengajaran dan pentaksiran dari perspektif pelaksanaan pentaksiran bilik darjah. [Curriculum, teaching, and assessment from the perspective of classroom assessment implementation]. *Asian People Journal, 3*(1), 152-161. https://doi.org/10.37231/apj.2020.3.1.175

Ary, D., Jacobs, L. C., Irvine, C. K. S., & Walker, D. (2019). *Introduction to research in education*. Wadsworth Cengage Learning.

Baidoo-Anu, D., Rasooli, A., DeLuca, C., & Cheng, L. (2023). Conceptions of classroom assessment and approaches to grading: Teachers' and students' perspectives. *Education Inquiry*. https://doi.org/10.1080/20004508.2023.2244136

Balbi, A., Bonilla, M. Otamendi, M. A., Curione, K., & Beltrán-Pellicer, P. (2020). Formative assessment and mathematics education: The perspective of in-service mathematics teachers. *Acta Scientiae, 24*(6), 236-269. https://doi.org/10.17648/acta.scientiae.7043

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *International Journal of Phytoremediation, 21*(1). https://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31. https://doi.org/10.1007/s11092-008-9068-5

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy and Practice, 25*(6), 551-575. https://doi.org/10.1080/0969594X.2018.1441807

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*(1), 8-21. https://doi.org/10.1177/003172170408600105

Boström, E., & Palm, T. (2023). The effect of a formative assessment practice on student achievement in mathematics. *Frontiers in Education, 8*. https://doi.org/10.3389/feduc.2023.1101192

Brandmo, C., Panadero, E., & Hopfenbeck, T. N. (2020). Bridging classroom assessment and self-regulated learning. *Assessment in Education: Principles, Policy & Practice, 27*(4), 319-331. https://doi.org/10.1080/0969594x.2020.1803589

Chua, Y. P. (2014). *Kaedah penyelidikan* [*Research method*]. McGraw Hill.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education*. Routledge. https://doi.org/10.4324/9781315456539

Fauziah, D., Mardiyana, & Saputro, D. R. S. (2018). Mathematics authentic assessment on statistics learning: The case for student mini projects. *Journal of Physics: Conference Series, 983*(1), 1-5. https://doi.org/10.1088/1742-6596/983/1/012123

Fjørtoft, H. (2020). Multimodal digital classroom assessments. *Computers & Education, 152*, Article 103892. https://doi.org/10.1016/j.compedu.2020.103892

Foster, C. (2022). Implementing confidence assessment in low-stakes, formative mathematics assessments. *International Journal of Science and Mathematics Education, 20*(7), 1411-1429. https://doi.org/10.1007/s10763-021-10207-9

Gezer, T., Wang, C., Polly, A., Martin, C., Pugalee, D., & Lambert, R. (2021). The relationship between formative assessment and summative assessment in primary grade mathematics classrooms. *International Electronic Journal of Elementary Education, 13*(5), 673-685. https://doi.org/10.26822/iejee.2021.220

Haj-Yahya, A., & Olsher, S. (2022). Preservice teachers' experiences with digital formative assessment in mathematics. *International Journal of Mathematical Education in Science and Technology, 53*(7), 1751-1769. https://doi.org/10.1080/0020739X.2020.1842527

Hashim, S., Zakariah, S. H., Taufek, F. A., Zulkifli, N. N., Che Lah, N. H., & Murniati, D. E. (2021). An observation on implementation of classroom assessment in technical and vocational education and training (TVET) subject area. *Journal of Technical Education and Training, 13*(3), 190-200. https://doi.org/10.30880/jtet.2021.13.03.019

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. https://doi.org/10.3102/003465430298487

Hidayat, R., Nugroho, I., Zainuddin, Z., & Ingai, T. A. (2023). A systematic review of analytical thinking skills in STEM education settings. *Information and Learning Sciences, 125*(7/8), 565-586. https://doi.org/10.1108/ils-06-2023-0070

Hidayat, R., Qi, T. Y., Ariffin, P. N. B. T., Hadzri, M. H. B. M., Chin, L. M., Ning, J. L. X., & Nasir, N. (2024). Online game-based learning in mathematics education among Generation Z: A systematic review. *International Electronic Journal of Mathematics Education, 19*(1), Article em0763. https://doi.org/10.29333/iejme/14024

Hidayat, R., Qudratuddarsi, H., Mazlan, N. H., & Mohd Zeki, M. Z. (2021). Evaluation of a test measuring mathematical modelling competency for Indonesian college students. *Journal of Nusantara Studies, 6*(2), 133-155. https://doi.org/10.24200/jonus.vol6iss2pp133-155

Hidayat, R., Syed Zamri, S. N. A., & Zulnaidi, H. (2018). does mastery of goal components mediate the relationship between metacognition and mathematical modelling competency? *Educational Sciences: Theory and Practice, 18*(3), 579-604. https://doi.org/10.12738/estp.2018.3.0108

Hidayat, R., Syed Zamri, S. N. A., Zulnaidi, H., & Yuanita, P. (2020). Meta-cognitive behaviour and mathematical modelling competency: Mediating effect of performance goals. *Heliyon, 6*(4), Article e03800. https://doi.org/10.1016/j.heliyon.2020.e03800

Joshi, D. R., & Rawal, M. (2021). Mathematics teachers standing on the utilization of digital resources in Kathmandu, Nepal. *Contemporary Mathematics and Science Education, 2*(1), Article ep21004. https://doi.org/10.30935/conmaths/9679

Joshi, D. R., Adhikari, K. P., Khanal, J., & Belbase, S. (2023). Impact of digital skills of mathematics teachers to promote students' communication behaviour in the classroom. *Contemporary Educational Technology, 15*(4), Article ep454. https://doi.org/10.30935/cedtech/13495

Kamarudin, N., AlRaqadi, Z., Alhunaini, S., & Zaremohzzabieh, Z. (2021). Assessment practices of mathematics teachers in Oman. *Turkish Journal of Computer and Mathematics Education, 12*(14), 4217-4224. https://doi.org/10.17762/turcomat.v12i14.11236

Konting, M. M. (2000). *Kaedah penyelidikan pendidikan* [*Educational research methods*]. Dewan Bahasa dan Pustaka [Language and Library Council].

Kultur, Y. Z., & Kutlu, M. O. (2021). The effect of formative assessment on high school students' mathematics achievement and attitudes. *Journal of Pedagogical Research, 5*(4), 155-171. https://doi.org/10.33902/JPR.2021474302

Landell, K. (1997). *Management by menu*. Wilay and Sms Inc.

Lestariani, I., Sujadi, I., & Pramudya, I. (2018). The implementation of portfolio assessment by the educators on the mathematics learning process in senior high school. *Journal of Physics: Conference Series, 1022*, Article 012011. https://doi.org/10.1088/1742-6596/1022/1/012011

Lund, A. M. (2001). Measuring usability with the use questionnaire. *Usability Interface, 8*(2), 3-6.

Maslan, M., & Mohd Nor, M. Y. (2020). Kebolehlaksanaan pentaksiran bilik darjah (PBD) secara atas talian sepanjang perintah kawalan pergerakan di Daerah Sentul, Kuala Lumpur [Feasibility of online classroom assessment (CA) during the movement control order in Sentul District, Kuala Lumpur]. *Webinar Seminar Nasional Pendidikan 2020* [*National Education Seminar Webinar 2020*], *1*(1), 213-218.

McMillan, J. H., & Moore, S. (2020). Better being wrong (sometimes): Classroom assessment that enhances student learning and motivation. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 93*(2), 85-92. https://doi.org/10.1080/00098655.2020.1721414

Ministry of Education Malaysia. (2015). *Dokumen standard kurikulum dan pentaksiran matematik tingkatan satu kurikulum standard sekolah menengah* [*Curriculum and assessment standard document mathematics form one standard middle school*]. KPM, Bahagian Pembangunan Kurikulum.

Ministry of Education Malaysia. (2019*). Panduan pelaksanaan pentaksiran bilik darjah* [*Classroom assessment implementation guide*]. KPM, Bahagian Pembangunan Kurikulum.

Mohd Isa, A., Mydin, A. A., Abdul Ghani Kanesan Abdullah, A. G., & Md Rasidi, W. F. (2020). Transformasi pendidikan tahap 1: Peperiksaan ke pentaksiran bilik darjah, kesan terhadap autonomi guru [Educational transformation level 1: Examination to classroom assessment, impact on teacher autonomy]. In N. F. Habidin, T. W. Tuan Chik, S. Y. Yee Ong, U. A. Muhammad, & N. Mohd Fuzi (Eds.), *Isu dan cabaran dalam pendidikan: Strategi dan inovasi* [*Issues and challenges in education: Strategies and innovations*]. Kaizentrenovation Sdn. Bhd.

Mohd Tajudin, N., Zamzamir, Z., Rahmat, F., Shafie, S., & Hussin, N.H. (2022). *Panduan pelaksanaan kajian reka bentuk dan pembangunan dan reka bentuk kajian pembangunan* [Design study implementation guide and development and design of developmental studies]. Universiti Pendidikan Sultan Idris.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. McGraw-Hill.

Reinholz, D. L. (2015). Peer-assisted reflection: A design-based intervention for improving success in calculus. *International Journal of Research in Undergraduate Mathematics Education, 1*, 234-267. https://doi.org/10.1007/s40753-015-0005-y

Rumanová, L., Vallo, D., & Záhorská, J. (2020). The impact of formative assessment on results of secondary school pupils in mathematics: One case of schools in Slovakia. *TEM Journal, 9*(3), 1200-1207. https://doi.org/10.18421/TEM93-47

Sanaeifar, S. H., & Mirshojaee, S. B. (2020). Optimizing students' engagement through implementing peer-assessment practice in iranian public high school: An action research. *Theory and Practice in Language Studies, 10*(8), 940-946. https://doi.org/10.17507/tpls.1008.12

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14. https://doi.org/10.3102/0013189X029007004

Siraj, S., Tony Lim Abdullah, M. R., & Muhammad Rozkee, R. (2022). *Pendekatan penyelidikan reka bentuk dan pembangunan* [*Design research and development approaches*]. Universiti Pendidikan Sultan Idris.

Staberg, R. L., Febri, M. I. M., Gjøvik, Ø., Sikko, S. A., & Pepin, B. (2022). Science teachers' interactions with resources for formative assessment purposes. *Educational Assessment, Evaluation and Accountability, 35*, 5-35. https://doi.org/10.1007/s11092-022-09401-2

Syaifuddin, M. (2019). The effect of students' perception on classroom assessment to students' attitudes. *Journal of Physics: Conference Series, 1280*, Article 042027. https://doi.org/10.1088/1742-6596/1280/4/042027

Taherdoost, H. (2016). Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. *International Journal of Academic Research in Management, 5*(3), 28-36. https://doi.org/10.2139/ssrn.3205040

Thinwiangthong, S., Eddy, C. M., & Inprasitha, M. (2020). Mathematics teachers abilities in developing formative assessment after the introduction of lesson study and open approach innovations. *Malaysian Journal of Learning and Instruction, 17*(1), 101-132. https://doi.org/10.32890/mjli2020.17.1.5

Veldhuis, M., & Van Den Heuvel-Panhuizen, M. (2019). Supporting primary school teachers' classroom assessment in mathematics education: Effects on student achievement. *Mathematics Education Research Journal, 32*(3), 449-471. https://doi.org/10.1007/s13394-019-00270-5

Wiliam, D., & Thompson, M. (2008). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53-82). Lawrence Erlbaum Associates. https://doi.org/10.4324/9781315086545-3

Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research, 25*(5), 508-518. https://doi.org/10.1177/0193945903252998

Zamri, N., & Hamzah, M. (2019). Teachers' competency in implementation of classroom assessment in learning. *Creative Education, 10*, 2939-2946. https://doi.org/10.4236/ce.2019.1012218

Zhao, X., Van Den Heuvel-Panhuizen, M., & Veldhuis, M. (2019). Insights Chinese primary mathematics teachers gained into their students' learning from using classroom assessment techniques. *Education Sciences, 9*(2), Article 150. https://doi.org/10.3390/educsci9020150

Zulliger, S., Buholzer, A., & Ruelmann, M. (2022). Observed quality of formative peer and self-assessment in everyday mathematics teaching and its effects on student performance. *European Journal of Educational Research, 11*(2), 663-680. https://doi.org/10.12973/eu-jer.11.2.663