REID (Research and Evaluation in Education), 10(1), 2024, 64-79
Available online at: http://journal.uny.ac.id/index.php/reid

Constructing and providing content validity evidence through the Aiken's V index based on the experts' judgments of the instrument to measure mathematical problem-solving skills

Nia Kania[1,2]; Yaya S. Kusumah*[1]; Jarnawi Afgani Dahlan[1]; Elah Nurlaelah[1]; Ferit Gürbüz[3]; Ebenezer Bonyah[4]
[1]Universitas Pendidikan Indonesia, Indonesia
[2]Universitas Majalengka, Indonesia
[3]Kırklareli University, Turkey
[4]Akenten Appiah Menka University of Skills Training and Entrepreneurial Development, Ghana
*Corresponding Author. E-mail: yskusumah@upi.edu

ARTICLE INFO

Article History
Submitted:
27 January 2024
Revised:
6 April 2024
Accepted:
14 June 2024

Keywords
assessment practices; classroom assessment; instruction; feedback; content validity

Scan Me:

ABSTRACT

Test content-based proof of validity is a type of evidence that supports the validity of a measuring instrument. This research aims to develop a mathematical problem-solving assessment instrument utilizing five experts. This study is classified as developmental research and follows a research design that includes two separate stages: the preliminary design stage and the prototype stage. However, its application is restricted to Prototype 1 and Prototype 2, specifically for expert evaluation. This instrument was designed explicitly for grade VIII students studying mathematics, covering all the topics from the odd semesters. The analysis progressed through three distinct stages—curriculum analysis, content analysis, and context analysis—each contributing to a comprehensive understanding of instructional resources. The study sought to narrow the gap between theoretical knowledge and practical application in mathematics education by incorporating real-world context. Surveys have revealed difficulties in answering mathematical problems, highlighting the need to address gaps in learning to improve competency. The careful and thorough construction of test instruments, considering factors such as validity, established the foundation for creating accurate assessment tools. The content validity assessment by the expert panel, with scores ranging from 0.817 to 0.884 based on the V-Aiken category, confirms that the instrument is vital in assessing students' mathematical problem-solving skills, and the implementation of this study yielded many valuable insights for educators and academics. This study helps improve mathematics education resources and evaluations to promote mathematical thinking.

To cite this article (in APA style):
Kania, N., Kusumah, Y., Dahlan, J., Nurlaelah, E., Gürbüz, F., & Bonyah, E. (2024). Constructing and providing content validity evidence through the Aiken's V index based on the experts' judgments of the instrument to measure mathematical problem-solving skills. *REID (Research and Evaluation in Education), 10*(1), 64-79. doi:https://doi.org/10.21831/reid.v10i1.71032

INTRODUCTION

The importance of problem-solving in mathematics education cannot be understated. Within mathematics education, mathematical problem-solving (MPS) skills promote analytical reasoning and improve students' capacity to address intricate problems. In the field of mathematics, MPS plays a crucial role in students' cognitive concepts (del Olmo-Muñoz et al., 2022; English, 2023), cognitive processes (Agustan et al., 2017; Haeruddin et al., 2020; Hollenstein et al., 2022), and acquisition of mathematical knowledge (Verschaffel et al., 2020, pp. 1-16) for effective goal achievement (Abdullah et al., 2015; Dewi & Kusumah, 2014).

MPS is commonly regarded as intricate and is recognized as a fundamental of mathematics education (Çekmez & Bülbül, 2018; Juniati & Budayasa, 2022; Pongsakdi, 2020). A study by Pambudi et al. (2020) reveals that every student encounters difficulties when learning mathematics in the classroom, which further supports this. When dealing with mathematical problems, students often encounter the following deficiencies: failure to properly understand the nature, context, and significance of a problem (Öztürk et al., 2020); failure to gather information from multiple sources (Ulitzsch et al., 2022); failure to evaluate methods and solutions after an incident (García et al., 2019); failure to apply problem-solving techniques effectively and efficiently (Jiang et al., 2021).

One factor contributing to the low MPS is the inadequacy of the practice questions in developing problem-solving skills. The questions given to students do not refer to practicing problem-solving (Risma & Yulkifli, 2021). This kind of situation causes students' problem-solving skills to be low (Purnomo et al., 2022). Teachers face the issue of developing suitable tools to assess the relevant skills. Training problem-solving skills can be challenging in education (Schöbel et al., 2023; Seepiwsiw & Seehamongkon, 2023). Contemporary mathematics teachers must create appropriate evaluation instruments (Ling & Loh, 2023; Makarova et al., 2021), utilizing various resources and practical scenarios (Apino & Retnawati, 2016). Net (2023), Boonen et al. (2016), and Hoseana (2024) argued that teachers have a crucial role in creating learning settings that foster the development of these vital skills.

Intensive practice is crucial in enhancing cognitive abilities, including mathematical problem-solving abilities. With lots of practice, students not only improve their aptitude for understanding mathematical principles (Lee & Ward-Penny, 2022) but also grow the capacity to utilise their MPS abilities in many situations efficiently (Hourigan & Leavy, 2023; Saadati et al., 2023). According to the research conducted by Malepa-Qhobela and Mosimege (2022) mathematical problem-solving questions aid in developing students' mathematical knowledge and enhance their capacity to tackle real-world situations, improving their generic problem-solving ability.

It is crucial to create a robust evaluation system that allows students to accurately express and demonstrate their academic achievements in line with the real-life conditions they meet during their learning process, including within mathematical education. The primary purpose of assessment is to provide educators with the information they need to make decisions based on student progress (Pastore, 2023) and adjust their teaching strategies wisely (Saadati & Celis, 2022). Apart from that, according to research by Jamil et al. (2023), evaluating cognitive abilities is considered one of the most important aspects of the learning process. Measuring cognitive abilities is critical in studying cognitive research (Ryoo et al., 2022). By evaluating cognitive abilities, educators can better understand the extent to which students can apply their knowledge and skills in everyday life and identify areas that require special attention in the teaching and learning process. Assessment is not just a tool for providing grades but is also a means of understanding the extent to which students understand the subject matter and their ability to apply it.

In educational settings, the assessment framework consists of three primary forms: assessment as learning (Kumar & Moral, 2022), assessment for learning (Yigletu et al., 2023), and assessment of learning (Plessis & Ewing, 2017). These many forms have distinct objectives for assessing students' knowledge, abilities, and development throughout their academic journey. Assessment as learning entails students actively participating in self-assessment and introspection to improve their comprehension and learning journey. Assessment for learning is centered around delivering feedback and assistance to students throughout their learning process, directing them toward enhancement and proficiency in the subject area. Finally, Spinney and Kerr (2023) state that the assessment of learning evaluates students' accomplishments and expertise in specific learning goals using diverse assessment techniques, such as tests, quizzes, or performance evaluations. These methods may encompass many assessment instruments, such as multiple-choice questions, essays, practical demonstrations, or other evaluating measures. The selection of the assessment method is contingent upon the educational goals, the content's characteristics, and the evaluation's intended results (Rao & Banerjee, 2023). Research from reliable sources, like Sutiarso et al. (2022) and Retnawati et al. (2018), shows the importance of using various evaluation methods to accurately measure students' learning outcomes and give them helpful feedback for ongoing improvement.

Several scholars, including Botelho et al. (2023), Desha et al. (2021a), and Fathiyah et al. (2019), have created diverse evaluation instruments that make substantial contributions in the field of assessment in mathematical problem-solving. Their endeavors to construct these assessment instruments demonstrate a profound fascination with quantifying performance and educational achievements. Botelho et al. (2023) introduce novel methods for creating assessment instruments that comprehensively overview students' academic performance. In contrast, the study by Desha et al. (2021b) highlights the importance of creating assessment instruments that can precisely gauge comprehension of crucial principles within a particular field. Furthermore, Fathiyah et al. (2019) have made a valuable contribution by introducing an assessment tool that considers psychometric factors to guarantee the accuracy and credibility of the measurements. By working together, academics have paved the way for the creation of advanced and contextually relevant assessment tools, enhancing evaluation procedures in diverse educational domains.

Although assessment is widely recognized for its advantages, it faces various obstacles that hinder its positive effect on the teaching and learning of mathematics. These issues include insufficient assessment resources, excessive instructor workload, and low student attendance at school (Buabeng et al., 2019). Studies have improved evaluation techniques, but semester-specific testing and development are scarce. Many scholars have developed evaluation frameworks for extended or comprehensive assessments, but none for semester-specific assessments. This research emphasizes the need for more significant research on one-term assessment approaches to evaluate student achievement fully. By continuing this research, scientists can better understand short-term assessment issues and build better methods.

A novelty measuring method was created and applied in this work to improve MPS evaluation in one semester. This endeavor involves mathematics content, educational mathematics, and measurement professionals to ensure tool validity. Diverse perspectives should help define MPS and ensure the instrument fits the assessment aims. The research tool should accurately assess students' MPS skills across one semester. This study suggests that mathematics evaluation approaches are getting more focused and relevant.

METHOD

This study is categorized as developmental research, using a research design consisting of two distinct stages: the preliminary design stage and the prototype stage (Kennedy-Clark, 2013; Plomp, 2013). The preliminary stage begins the research process by collecting students' perspectives and aspirations based on their experiences studying mathematics. During this stage, the highlighted shortcomings are thoroughly examined, and a detailed strategy is developed to solve these problems in the product design. Diverse research methods, including examinations, interviews, and questionnaires, are used to comprehensively comprehend student requirements, establishing a robust basis for producing successful and relevant products. The prototyping phase concurrently enables problem-solving by systematically revising the product design and engaging researchers, professionals, students, and instructors to guarantee its adaptability to real-world requirements. This phase includes developing mathematical problem-solving tools, expert evaluation, and controlled trials involving students with varying abilities. Researchers develop and describe preliminary designs and concepts for the eventual product during this stage. Nevertheless, to maintain confidentiality, participants were anonymous, and the collected data were used only for research purposes.

As presented in Figure 1, the assessment procedure in the prototyping phase starts with Prototype 1, which must undergo trials to ascertain the efficacy of its design. Subsequently, Prototype 2 received a rigorous evaluation by domain specialists, particularly those specializing in mathematics education. Prototype 3 was subjected to a restricted testing phase, including students, which allowed the detection of technological vulnerabilities and the identification of required enhancements. Prototype 4 ultimately undertook a comprehensive four-week field test, integrating the viewpoints of Mathematics educators. The outcomes of this extensive phase act as the foundation for ongoing enhancement and advancement of the product, guaranteeing its compatibility with the changing requirements of students. This study focuses on comprehending the development model at the Prototype 2 phase to establish validity and dependability via expert evaluation. The latter two phases of Prototype 2 are crucial for further investigation as well as verification.

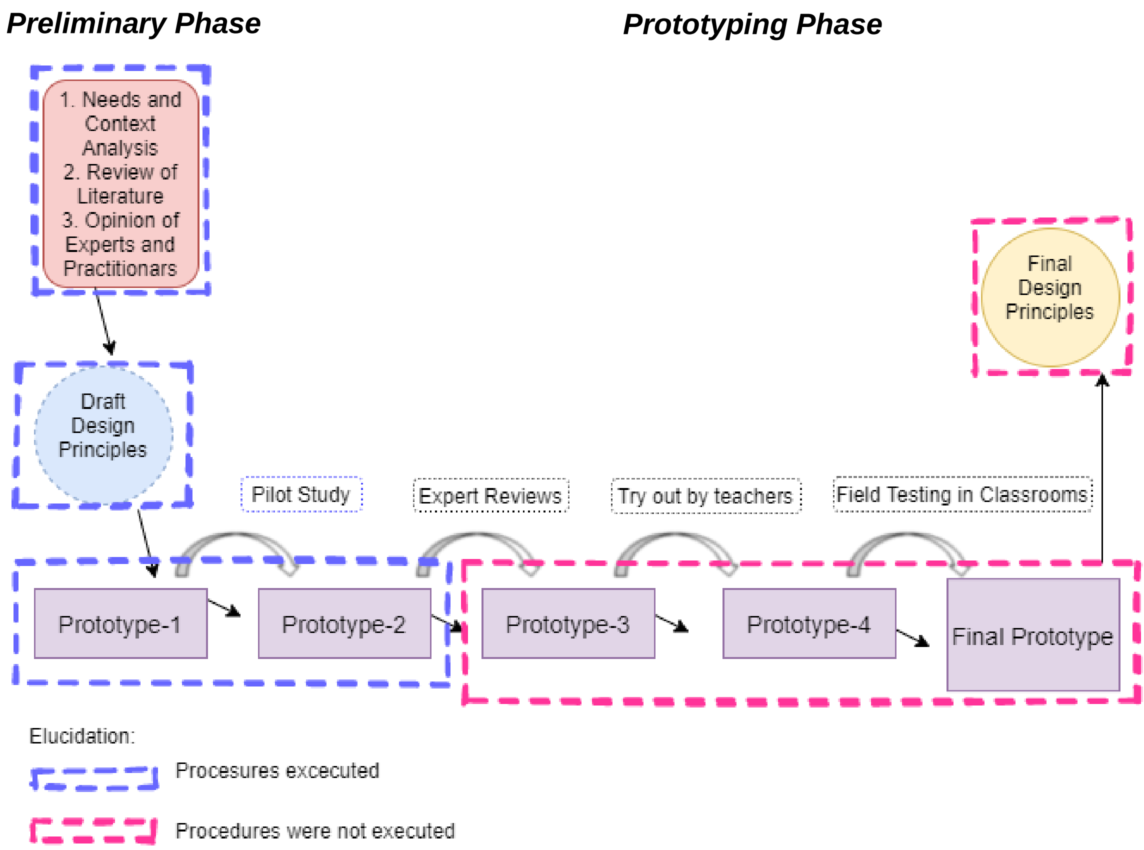

Figure 1. Revised Diagram of the Development Chart

Before commencing the research, researchers undertook a crucial step by constructing indicators to quantitatively assess proficiency in solving mathematical problems. This stage is crucial as it offers precise guidance in assessing the study variables that need to be accomplished. By establishing suitable indicators, researchers can construct a robust framework to effectively guide their research process, guaranteeing the acquired results' relevance. Furthermore, defined metrics also assist researchers in quantifying progress and accomplishments across the many stages of their research, enabling them to draw more insightful and precise conclusions regarding the issues they investigate. The instrument devised in this study aims to cater to pupils in the eighth grade of junior high school. All the selected coursework consists of odd semester content condensed into one semester.

Validity refers to the degree of precision and accuracy a measuring device shows when performing its intended purpose. The study conducted a content validity assessment, which evaluates the correctness of a test instrument about its content (Latisma, 2015). The content validity test in this study was administered by a panel of five specialists, including two experts in mathematics learning, two experts in school mathematics materials, and one expert in measurement, all with a specialization in mathematics education. A panel of five validators evaluated a content validation sheet for MPS problems, the study tool. The evaluation was conducted by assigning a numerical rating on a five-point scale. Equation (1) was utilized to assess the validity of test findings for experts and users using the Aiken methodology, in which V = item validity index, *r* = score of the rater's chosen category, l_o = lowest score in the scoring category, *c* = category that the rater can choose, and *n* = number of raters (Retnawati, 2014).

$$V = \frac{\sum s}{n(c-1)} \quad (1)$$

Aiken's rules offer a valuable initial framework for evaluating content validity using Aiken's V coefficient. Researchers evaluated the method and recommendations by Retnawati (2016) while incorporating additional metrics or criteria to provide a thorough and robust evaluation of content validity. Equation (1) indicates that a V-Aiken value less than 0.600 is classified as "poor" and a score between 0.600 and 0.880 is classified as "good". A result over 0.800 is classified as "excellent". Retnawati (2016) states that an Aiken value index below 0.4 indicates poor validity, while a range of 0.4 to 0.8 suggests moderate validity, and a value over 0.8 indicates high validity.

Research of a qualitative nature was carried out to assess the readability of the material. The evaluation included the use of language, writing strategies, punctuation application, font choice, image incorporation, and sentence duration. The assessment involved the use of a questionnaire with statements and five alternative answers. The selections were ranked in the sequence of 5 for excellent, 4 for acceptable, 3 for average, 2 for below-average, and 1 for unacceptable.

FINDINGS AND DISCUSSION

Notable researchers in the field of mathematical problem-solving include computer scientists like Turing (1954), renowned for his significant contributions to computing and algorithmic problem-solving. In addition, mathematicians like Polya (1973) have gained recognition for their contributions to developing mathematical problem-solving procedures, notably through Pólya's renowned book "How to Solve It." Psychologists like Piaget and Vygotsky (1994) have significantly contributed to our knowledge of how humans learn and solve mathematics issues using cognitive development approaches. Our comprehension of solving mathematical problems constantly expands by conducting research and making valuable contributions.

Table 1. Dimension and Indicator of MPS

Aspect	Definition	Indicator
Cognitive Processes	Information collection, viewpoint exploration, and a precise grasp of the problem's nature, context, and ramifications are necessary to understand a challenge.	The capacity to clearly and precisely define issues
The Gathering of Data	Data collection skills entail gathering information from a range of sources.	The capacity to plan or strategize by thinking through the required actions.
Efficient Execution	Capacity to implement strategies with effectiveness and efficiency to solve issues	The capacity to assess plan implementation outcomes to pinpoint areas for possible additions or improvements.
Reflective Evaluation	Capacity to assess methods and solutions after the fact	The capacity to evaluate the efficacy and efficiency of solutions or processes by contrasting them with benchmarks or other options.
Analytical problem-solving	The capacity to control complexity—to tackle complicated issues by breaking them down into more manageable, easily understood components	The capacity to create or formulate exact, in-depth solutions for every aspect of the solved problem.

This study examines many mathematical problem-solving indicators. Understanding and applying mathematics to problems is crucial. Detecting key facts, formulating problems clearly, and choosing solutions are evidence of problem-solving abilities. The ability to apply and evaluate strategies is a problem-solving strategy indication. Finally, mathematical communication skills involve organizing solutions, using appropriate terminology, and communicating rationally. This research examines these qualities and indicators to understand and enhance mathematical problem-solving. The indicators and dimensions used in this study are based on a synthesis of many experts in problem-solving. The synthesis process was conducted utilizing the Atlas.ti program to generate the aspects, definitions, and indicators as presented in Table 1.

By synthesizing several expert viewpoints on the subject, it is anticipated that various dimensions and indications about the resolution of mathematical issues can be developed (Ramdani et al., 2019; Datt, 2021). This dimension encompasses multiple crucial elements in mathematical problem-solving, including comprehension of mathematical concepts, the capacity to articulate and resolve problems, proficiency in employing efficient problem-solving strategies, and the aptitude to communicate solutions with clarity and precision. Indicators about this dimension enable researchers to assess an individual's proficiency in solving mathematical issues. Researchers can create comprehensive evaluation instruments to measure mathematical problem-solving abilities by comprehending these dimensions and signs.

Our study primarily centered on content structure analysis, using a methodology that included surveys of students and instructors and expert evaluations to determine content validity. This analytical approach provides a profound understanding of the perspectives and opinions of students and instructors, enabling a comprehensive comprehension of the overall structure of the content. These surveys provide firsthand perspectives from essential participants in the learning process, while professional evaluations help guarantee the validity of the material used. This introduction will provide a thorough overview of our research methodology to analyze content structure, examine data from direct participants, and maintain the ongoing validity of ideas via expert perspectives.

Preliminary Phase

The preliminary phase encompasses the needs and context analyses as well as the review of litera-ture and opinions of experts and practitioners.

Needs and Context Analysis

This study included administering questionnaires to students to ascertain their comprehension of the idea and gauge their level of interest in and attitudes toward the material studied. The current phase is built upon the findings derived from prior studies, which have been recorded in several journals. The analysis found that pupils had difficulties while attempting to solve non-routine situations. The crux of the issue seems to be in the student's inclination to commit notions and theories to memory without cultivating a profound comprehension. Moreover, the absence of regular experience in addressing topics that require critical thinking further intensifies this issue. Consequently, students lack the skills to effectively use their knowledge of scientific facts and approach scientific problem-solving. These results emphasize the need to tackle learning disparities to enhance students' competence in mathematical problem-solving skills.

In addition to administering questionnaires to students, this study also issued questionnaires to two mathematics educators to ascertain their viewpoints on comprehending ideas and assess students' mathematical problem-solving abilities. The analysis found that pupils had difficulties while attempting to solve non-routine situations. The fundamental issue is the kids' limited proficiency in mathematics from the beginning. Furthermore, the absence of regular experience in addressing topics that require critical thinking further intensifies this issue. These results emphasize the significance of addressing learning disparities to enhance students' ability to cope and improve their mathematical problem-solving skills.

Review of Literature

Subsequently, researchers conducted a literature review to cater to the requirements and analyse the content. The objective is to identify gaps in knowledge. Researchers can discern areas of knowledge that require further investigation or resolution by compiling data from current literature. It is crucial to establish the direction and focus of the needs and content analysis that will be conducted. Additionally, it is necessary to support material analysis and development plans. The data from literature research can be utilized to build suitable analysis methods and create materials that fulfill specific requirements. This entails the creation of instructional materials, educational approaches, or actionable suggestions derived from pertinent literature discoveries.

Opinions of Experts and Practitioners

The concluding phase of the preliminary phase, specifically the input from experts and practitioners, entails engaging in dialogue with the promoter team and educators in the relevant field. This aims to gather information directly from experts and practitioners with extensive experience and a thorough understanding of the study topic or project. The aims of this message encompass, *firstly*, the verification of concepts and plans. Obtaining advice from experts and practitioners is beneficial for verifying the thoughts and strategies developed in the previous step. They can offer valuable insights into the actual requirements, potential obstacles, and the practicality and significance of suggested remedies. *Secondly*, it acquires a pragmatic viewpoint. Experts in the field have firsthand knowledge of pertinent circumstances and contexts. Their feedback can offer a pragmatic and tangible viewpoint on the solution or plan, ensuring its effective implementation. *Thirdly*, it identifies opportunities and potential for innovation. By engaging with experts and practitioners, researchers can discover novel prospects or untapped innovation potential that may have been previously overlooked. Their contribution has the potential to stimulate novel ideas or guide research toward more efficient and pertinent paths.

The analysis results obtained during the preliminary phase serve as a reference for developing or creating problem-solving tools. The objective is to guarantee that the developed instrument is robust and applicable to assessing MPS. Moreover, examining outcomes can guide the creation of more efficient teaching and learning methods. One can improve and refine the instrument by incorporating feedback from students and teachers to enhance the efficacy of measuring problem-solving skills. It is crucial to ensure that measurement findings yield precise and pertinent information to facilitate informed decision-making in the educational setting.

Draft Design Principles

Content structure analysis comprises three primary stages: curriculum, content, and context analysis. The first phase, curriculum analysis, specifically targeted the educational content designed for eighth-grade pupils during the first half of the academic year, following Curriculum 2013 guidelines. The five resources identified include number patterns, Cartesian coordinates, relations and functions, straight-line equations, and systems of linear equations in two variables. The research yielded competency achievement indicators (GPA) based on *Kompetensi Dasar* (KD) or the basic competency for each subject. These indicators were combined with mathematical problem-solving indicators to provide five interconnected indicators.

The primary dimensions of this study are the outcome of synthesizing several specialists' perspectives on problem-solving skills. They offer their perspectives and knowledge to create a solid conceptual base. The outcomes of this collaboration include a thorough comprehension of numerous vital facets of issue resolution. These dimensions serve as a solid basis for in-depth data analysis and interpretation and reflect the conceptual framework. Consequently, the findings of this study not only provide a thorough comprehension of issue-solving but also pave the way for additional research and the creation of novel ideas in this area.

The second phase, known as content analysis, involves using middle school mathematics textbooks as a resource to examine and evaluate the measurement of MPS and the organization of mathematical content, as determined by specialists in the field. This analysis aims to assure compliance with the intended level of comprehension of mathematical ideas.

The last phase, context analysis, aims to relate the five chosen resources' context to daily life challenges or events. The study relies on journals, scholarly publications, and chemical books as sources. Incorporating context into mathematics education aims to enable students to comprehend and use mathematical principles in practical scenarios, thus establishing a link between mathematical learning and real-world applications.

Prototype-1: Pilot Study

Instrument preparation involved using a question grid and emphasizing mathematical problem-solving characteristics as crucial parameters for designing questions. The findings from analyzing the curriculum, content, and context, as well as the results of surveys conducted with students and teachers, serve as a basis for creating the question blueprint. The question blueprint comprises KD, GPA, question indications, and mathematical problem-solving features. Subsequently, the questions are constructed using the blueprint and integrating mathematical problem-solving elements.

Prototype-2: Expert Reviews

A panel of five experts in the relevant field conducted the content validity assessment. The content validity test in this research was conducted by a panel of five professionals, consisting of two specialists in mathematics learning, two specialists in school mathematics resources, and one specialist in measurement, all with expertise in mathematics education. The mathematical problem-solving curriculum in the first semester of middle school class VIII comprises five topics: number patterns, Cartesian coordinates, relations and functions, straight-line equations, and systems of linear equations in two variables. The content validity evaluation instrument for mathematical problem-solving problems consists of material/content, construction, and language. In all, there are 23 sub-aspects. The outcomes of the content validity evaluation serve as a reference for rectifying any instrument questions that remain erroneous.

When constructing test instruments, it is necessary to consider four fundamental concepts: validity, reliability, objectivity, and norms (Guzmán et al., 2021). Validity refers to an instrument's capability to measure the intended quantity accurately. In contrast, dependability refers to an instrument that consistently produces the same data when used multiple times to measure the same object. Instrument validation is an essential step in developing and evaluating instruments. Validity refers to the degree of accuracy in testing the components and assessing their suitability for interpretation. The validation process involves collecting evidence to establish a scientific foundation for interpreting the scores, as outlined in the intended use assessment tool. Score evaluation results can be interpreted based on the tool's intended purpose. Nevertheless, the initial procedure to obtain an accurate interpretation is to validate the instrument beforehand. According to the validator's evaluation of the measurement tools for mathematical problem-solving, each indicator is displayed in Table 2.

The absence of invalid items in evaluating mathematical problem-solving (MPS) measurement instruments using V-Aiken indicates a strong and dependable assessment procedure. This result highlights the efficacy of the instruments in precisely assessing the specific concepts within each MPS category. The instruments' strong validity scores, which range from 0.817 to 0.884 across several categories, demonstrate their capacity to measure students' proficiency in mathematical problem-solving skills accurately. The positive validation confirms the instruments' correctness and strengthens their legitimacy and usefulness in academic assessments. Educators and researchers can use these measurement tools to gather precise and essential information about students' mathematical problem-solving abilities. This data can then be used to enhance instructional strategies and develop curriculum in mathematics education.

Table 2. Validity of Test Items based on *V-Aiken*

Material	Question Number	Indicator MPS	Validator					∑s	n(c-1)	V	Categories
			1	2	3	4	5				
Number Pattern	10	MPS1	4	5	4	4	4	17	20	0.832609	Valid
	11	MPS2	4	5	4	4	4	17	20	0.847826	Valid
	14	MPS2	4	4	4	4	4	16	20	0.817391	Valid
	15	MPS1	4	5	4	4	5	17	20	0.856522	Valid
	16	MPS3	5	5	5	4	4	17	20	0.863043	Valid
	17	MPS2	4	5	4	4	4	17	20	0.843478	Valid
	18	MPS4	4	4	4	4	4	16	20	0.821739	Valid
	19	MPS5	4	5	4	4	4	17	20	0.845652	Valid
Cartesian Coordinates	1	MPS1	5	5	5	4	4	17	20	0.865217	Valid
	2	MPS2	5	5	5	4	4	18	20	0.884783	Valid
	9	MPS4	4	5	5	4	4	17	20	0.85	Valid
	12	MPS3	5	4	5	4	4	17	20	0.858696	Valid
	13	MPS2	4	4	4	5	4	17	20	0.856522	Valid
	14	MPS2	4	5	4	4	4	17	20	0.832609	Valid
	19	MPS5	4	4	5	4	4	17	20	0.854348	Valid
	20	MPS5	4	4	4	4	4	16	20	0.823913	Valid
	21	MPS5	4	4	4	4	4	17	20	0.841304	Valid
	23	MPS3	4	4	4	4	4	17	20	0.854348	Valid
Relationships and Functions	9	MPS1	4	5	4	4	4	17	20	0.836957	Valid
	10	MPS4	5	5	4	4	4	17	20	0.85	Valid
	12	MPS1	4	4	4	4	4	17	20	0.836957	Valid
	17	MPS4	5	5	4	4	4	17	20	0.867391	Valid
	18	MPS2	4	4	4	4	4	17	20	0.830435	Valid
	19	MPS4	4	4	5	4	4	17	20	0.856522	Valid
	23	MPS4	4	4	4	4	4	17	20	0.85	Valid
	24	MPS1	4	4	4	4	4	17	20	0.836957	Valid
	26	MPS5	4	5	5	4	4	17	20	0.863043	Valid
	27	MPS4	4	5	4	4	4	17	20	0.832609	Valid
	28	MPS3	4	5	5	4	4	17	20	0.856522	Valid
Straight Line Equation	6	MPS4	4	5	5	4	4	18	20	0.876087	Valid
	8	MPS2	4	4	4	4	4	17	20	0.847826	Valid
	9	MPS1	4	5	5	4	4	17	20	0.86087	Valid
	10	MPS5	4	4	4	4	4	17	20	0.85	Valid
	16	MPS1	4	5	4	4	4	17	20	0.83913	Valid
	17	MPS2	4	5	5	4	4	17	20	0.863043	Valid
	18	MPS5	4	4	5	4	4	17	20	0.847826	Valid
	19	MPS3	4	5	5	4	4	17	20	0.856522	Valid
	20	MPS2	4	5	4	4	4	17	20	0.841304	Valid
	25	MPS5	4	4	4	4	4	17	20	0.841304	Valid
	26	MPS4	4	5	4	4	4	17	20	0.871739	Valid
	28	MPS5	4	4	4	4	4	17	20	0.83913	Valid
	29	MPS3	5	4	4	4	4	17	20	0.836957	Valid
System of Two-Variable Linear Equations	3	MPS1	4	5	4	4	4	17	20	0.854348	Valid
	4	MPS2	5	5	4	4	4	17	20	0.850000	Valid
	9	MPS4	4	5	4	4	4	17	20	0.850000	Valid
	17	MPS5	4	4	4	4	4	17	20	0.83913	Valid
	19	MPS4	5	4	5	4	4	17	20	0.863043	Valid
	21	MPS3	4	5	5	4	4	17	20	0.854348	Valid

The findings on content validity were analyzed using Aiken's V. Table 3 displays the outcomes of the content validity based on assessed aspects. The utilization of V-Aiken for content validity analysis, as depicted in Table 3, offers a thorough review of diverse facets of assessing educational materials. The table classifies the discoveries according to many factors, such as material, construction, and language. Within the material aspect, the evaluation includes assessing factors such as the clarity of core competencies (*kompetensi inti* (KI))/basic competencies (*kompetensi dasar* (KD)) learning, appropriateness of purpose and material, alignment of the problem with the indicator, coherence between essential competencies and objectives, material, and questions, relevance of questions to the subject matter being studied, and the availability of a scoring rubric. Each sub-aspect is given a validity score ranging from 0.750 to 0.872, indicating different levels of validity. The material component demonstrates overall solid validity.

Table 3. Validity based on Assessed Aspects

Aspects	Sub Aspect	Expert					∑s	n(c-1)	V	Category
		1	2	3	4	5				
Material	Clarity of KI/KD learning	4	5	5	4	4	17	20	0.859	High
	Suitability of purpose and material	4	5	5	5	4	17	20	0.872	High
	Compatibility of the problem with the indicator	4	4	4	4	4	15	20	0.750	Medium
	Consistency between essential competencies with objectives, materials, and questions	4	4	4	5	4	17	20	0.865	High
	Suitability of the question with the subject matter studied	4	4	5	5	4	17	20	0.867	High
	The instrument is equipped with a scoring rubric	4	5	4	3	4	15	20	0.773	Medium
	Suitability of questions to the cognitive level of student ability	4	4	5	5	5	17	20	0.872	High
	Stimulate students to think more operationally	4	4	4	5	4	17	20	0.869	High
Construction	Clarity of the content of the material	4	5	5	5	4	17	20	0.874468	High
	Demands on the content of the material	4	5	4	5	4	17	20	0.860638	High
	The correctness of the content of the material	5	5	4	4	4	17	20	0.867021	High
	Ease of materials to understand	4	4	4	4	4	15	20	0.755319	Medium
	Suitability of the image to the content of the material	4	5	5	5	4	18	20	0.879787	High
	Clear scoring guidelines	4	5	4	4	4	17	20	0.86383	High
	The question points are formulated briefly and clearly	4	4	4	5	5	17	20	0.865957	High
	Clarity of instructions for working on the problem	5	5	4	4	4	17	20	0.868085	High
	The demands of the questions presented	4	4	4	4	5	17	20	0.857447	High
Language	Consistency in using terms, symbols, and units	4	5	4	4	4	17	20	0.86383	High
	Question items do not cause double interpretation	4	5	4	4	5	17	20	0.869149	High
	The use of standard Indonesian in question items	4	5	4	4	5	17	20	0.87234	High
	The composed language has subjects, predicates, objects, and captions	4	5	4	4	4	17	20	0.869149	High
	Sentences composed using effective sentences	4	5	4	3	4	15	20	0.77234	Medium
	Languages used according to EYD	4	5	4	5	4	17	20	0.867021	High

In the same way, the construction aspect assesses many variables, including demands, correctness, and ease of understanding of the material. These factors have received high validity scores, further confirming the quality of the instructional content. Finally, the language element evaluates the precision of language usage, avoidance of ambiguity, adherence to standard Indonesian, effective sentence structure, and conformance to language standards (*ejaan yang disempurnakan*/EYD). The consistently high validity scores across several language-related criteria confirm the linguistic excellence and appropriateness of the teaching materials. In summary, the comprehensive examination of content validity highlights the meticulousness and dependability of the assessment procedure, offering significant perspectives for educators, curriculum creators, and researchers to improve the calibre of educational materials.

Compatibility of the Instrument with the Curriculum

Curriculum revisions in Indonesia have taken place on multiple occasions, and throughout these revisions, assessment standards have consistently remained a crucial component of the curriculum. The National Education Standard serves as a fundamental framework for designing, executing, and overseeing education to achieve high-quality national education (Tahili et al., 2021). Instrument development is necessary regardless of the curriculum.

When assessing a program, it is necessary to establish the instrument's validity to verify its relevance and suitability in measuring the desired competencies. The validity of an instrument indicates its capacity to measure thinking skills accurately (Danczak et al., 2020; Tao et al., 2023; Wihardjo et al., 2023), specifically problem-solving skills. Content validity, when proven, instills confidence in the instrument's ability to deliver accurate and valuable information for evaluating and developing student skills. According to the professional opinions of Imran et al. (2024), Kim et al. (2024), and Zeighami et al. (2024), content validity refers to the capacity of a precise tool to assess students' abilities.

Nevertheless, this research is subject to certain limitations and suggests areas for future investigation. While the concept and validity results yield good indicators, this study is subject to various constraints. For instance, the constraints of a small sample size limit the ability to capture the full scope of the learning environment, and external variables may impact the accuracy of the measurements. Potential avenues for future research may involve advancing more intricate instruments, conducting trials on a broader scope, and further corroborating the findings through other pertinent methodologies.

However, this research must be continued to fulfill the requirements stated by Guzmán et al. (2021), who indicated that when designing test instruments, validity, reliability, objectivity, and norms should be considered. This research should continue to assess reliability. This study examines many mathematical problem-solving indicators. Understanding and applying maths to problems is crucial. Detecting key facts, formulating problems clearly, and choosing solutions are evidence of problem-solving abilities. The ability to apply and evaluate strategies is a problem-solving strategy indication. Finally, MPS involves organizing solutions, using appropriate mathematical terminology, and communicating rationally. This research examines these qualities and indicators to understand and enhance MPS.

CONCLUSION

Our study aimed to thoroughly investigate the content structure and validity by analyzing student and instructor surveys and expert reviews. This complex tapestry explored not only the distinct viewpoints of students and instructors but also guaranteed a comprehensive comprehension of the general structure of the topic. The primary accounts given by those who participated in the learning process and the assessments made by experts produced a harmonious collection of viewpoints that increased the study's credibility. The content structure analysis was carried out in three stages—curriculum analysis, content analysis, and context analysis—each leading to a deeper understanding of educational resources. Our work aims to establish a strong connection between abstract mathematical theories and practical applications by incorporating real-world examples into mathematics instruction. The surveys revealed difficulties in answering mathematical problems, emphasizing the need to address discrepancies in learning to improve competency. The careful and thorough development of test instruments, considering factors such as validity, lays the foundation for creating accurate and efficient assessment tools. The expert panel's content validity evaluation confirmed the potency and dependability of our measurement tools. These results demonstrate the accuracy of evaluating students' mathematical problem-solving abilities and strengthen the credibility of the instruments in academic assessments. Our study presents a wealth of valuable knowledge, encouraging educators and researchers to pursue the improvement of mathematics education materials and evaluations, leading to a new era of high-quality education.

ACKNOWLEDGMENTS

The author extended their gratitude to Lembaga Pengelola Dana Pendidikan (LPDP) for the invaluable assistance the first author received when finalizing the findings of this research.

DISCLOSURE STATEMENT

No conflicts of interest are linked to this publication. The corresponding author affirms that the paper has been reviewed and authorized for submission by all listed authors.

REFERENCES

Abdullah, A. H., Abidin, N. L. Z., & Ali, M. (2015). Analysis of students' errors in solving Higher Order Thinking Skills (HOTS) problems for the topic of fraction. *Asian Social Science*, *11*(21), 133–142. https://doi.org/10.5539/ass.v11n21p133

Agustan, S., Juniati, D., & Siswono, T. Y. E. (2017). *Investigating and analyzing prospective teacher's reflective thinking in solving mathematical problem: A case study of female-field dependent (FD) prospective teacher*. 040018. https://doi.org/10.1063/1.4983956

Apino, E., & Retnawati, H. (2016). Developing instructional design to improve mathematical higher order thinking skills of students. *Journal of Physics: Conference Series*, *755*(1), 011001. https://doi.org/10.1088/1742-6596/755/1/011001

Boonen, A. J. H., Reed, H. C., Schoonenboom, J., & Jolles, J. (2016). It's not a math lesson-we're learning to draw! teachers' use of visual representations in instructing word problem solving in sixth grade of elementary school. *Frontline Learning Research*, *4*(5), 34–61. https://doi.org/10.14786/flr.v4i5.245

Botelho, A., Baral, S., Erickson, J. A., Benachamardi, P., & Heffernan, N. T. (2023). Leveraging natural language processing to support automated assessment and feedback for student open responses in mathematics. *Journal of Computer Assisted Learning*, *39*(3), 823–840. https://doi.org/10.1111/jcal.12793

Buabeng, İ., Antingane, A. B., & Amoako, İ. (2019). Practices, challenges and perceived influence of classroom assessment on mathematics instruction. *International Journal of Assessment Tools in Education*, *6*(3), 476–486. https://doi.org/10.21449/ijate.616617

Çekmez, E., & Bülbül, B. Ö. (2018). An example of the use of dynamic mathematics software to create problem-solving environments that serve multiple purposes. *Interactive Learning Environments*, *26*(5), 654–663. https://doi.org/10.1080/10494820.2017.1385029

Danczak, S. M., Thompson, C. D., & Overton, T. L. (2020). Development and validation of an instrument to measure undergraduate chemistry students' critical thinking skills. *Chemistry Education Research and Practice*, *21*(1), 62–78. https://doi.org/10.1039/C8RP00130H

Datt, S. (2021). Theoretical basis for evolving taxonomy of the affective domain. *I-Manager's Journal on Educational Psychology*, *15*(2), 10-18. https://doi.org/10.26634/jpsy.15.2.17982 del Olmo-Muñoz, J., González-Calero, J. A., Diago, P. D., Arnau, D., & Arevalillo-Herráez, M. (2022). Using intra-task flexibility on an intelligent tutoring system to promote arithmetic problem-solving proficiency. *British Journal of Educational Technology*, *53*(6), 1976–1992. https://doi.org/10.1111/bjet.13228

Desha, C., Caldera, S., & Hutchinson, D. (2021a). Exploring the development of context appreciation in coursework that targets problem-solving for sustainable development. *International Journal of Sustainability in Higher Education*, *22*(5), 1186–1224. https://doi.org/10.1108/IJSHE-01-2020-0024

Desha, C., Caldera, S., & Hutchinson, D. (2021b). Exploring the development of context appreciation in coursework that targets problem-solving for sustainable development. *International Journal of Sustainability in Higher Education*, *22*(5), 1186–1224. https://doi.org/10.1108/IJSHE-01-2020-0024

Dewi, N. R., & Kusumah, Y. S. (2014). Developing test of high order mathematical thinking ability in integral calculus subject. *International Journal of Education and Research*, *2*(12), 101–108.

English, L. D. (2023). Ways of thinking in STEM-based problem solving. *ZDM – Mathematics Education*, *55*(7), 1219–1230. https://doi.org/10.1007/s11858-023-01474-7

Fathiyah, K. N., Alsa, A., & Setiyawati, D. (2019). Psychometric characteristic of positive affect scale within the academic setting. *REID (Research and Evaluation in Education)*, *5*(2), 120–129. https://doi.org/10.21831/reid.v5i2.25992

García, T., Boom, J., Kroesbergen, E. H., Núñez, J. C., & Rodríguez, C. (2019). Planning, execution, and revision in mathematics problem solving: Does the order of the phases matter? *Studies in Educational Evaluation*, *61*, 83–93. https://doi.org/10.1016/j.stueduc.2019.03.001

Guzmán, B., Rodríguez, C., Ferreira, R. A., & Hernández-Cabrera, J. A. (2021). Psychometric properties of the revised child mathematics anxiety questionnaire (CMAQ-R) for Spanish speaking children. *Psicologia Educativa*, *27*(2), 115–122. https://doi.org/10.5093/psed2020a17

Haeruddin, Prasetyo, Z. K., & Supahar. (2020). The development of a metacognition instrument for college students to solve physics problems. *International Journal of Instruction*, *13*(1), 767–782. https://doi.org/10.29333/iji.2020.13149a

Hollenstein, L., Thurnheer, S., & Vogt, F. (2022). Problem solving and digital transformation: acquiring skills through pretend play in kindergarten. *Education Sciences*, *12*(2), 92. https://doi.org/10.3390/educsci12020092

Hoseana, J. (2024). Using Gardner's three-squares problem for a group project in a mathematical problem solving module. *The Mathematics Enthusiast*, *21*(1–2), 467–484. https://doi.org/10.54870/1551-3440.1638

Hourigan, M., & Leavy, A. M. (2023). Elementary teachers' experience of engaging with teaching through problem solving using lesson study. *Mathematics Education Research Journal*, *35*(4), 901–927. https://doi.org/10.1007/s13394-022-00418-w

Imran, S., Yasmeen, R., & Mansoor, M. (2024). Development and validation of self-assessment instrument to measure the digital professionalism of healthcare professionals using social media. *BMC Medical Education*, *24*(1), 243. https://doi.org/10.1186/s12909-024-05142-6

Jamil, N., Belkacem, A.N. & Lakas, A. (2023). On enhancing students' cognitive abilities in online learning using brain activity and eye movements. *Education and Information Technologies,* 28, 4363–4397. https://doi.org/10.1007/s10639-022-11372-2

Jiang, R., Liu, R. de, Star, J., Zhen, R., Wang, J., Hong, W., Jiang, S., Sun, Y., & Fu, X. (2021).

How mathematics anxiety affects students' inflexible perseverance in mathematics problem-solving: Examining the mediating role of cognitive reflection. *British Journal of Educational Psychology*, *91*(1), 237–260. https://doi.org/10.1111/bjep.12364

Juniati, D., & Budayasa, I. K. (2022). The influence of cognitive and affective factors on the performance of prospective mathematics teachers. *European Journal of Educational Research*, *11*(3), 1379–1391. https://doi.org/10.12973/eu-jer.11.3.1379

Kennedy-Clark, S. (2013). Reflection: Research by design: Design-based research and the higher degree research student. *Journal of Learning Design*, *6*(2), 26–32. https://doi.org/10.5204/jld.v8i3.257

Kim, D., Ryoo, D., & Lee, M. (2024). Development of a scale to measure the parental competency of science-gifted students in South Korea. *Journal of Advanced Academics*, *35*(1), 125–155. https://doi.org/10.1177/1932202X231219277

Kumar, S. R., & Moral, C. S. (2022). Formative assessment tools for effective classroom. *I-Manager's Journal on School Educational Technology*, *17*(4), 1-10. https://doi.org/10.26634/jsch.17.4.18926

Latisma, D. J. (2015). Design and implementation of chemistry triangle oriented learning media on hydrocarbons. *The International Conference on Mathematics, Science, Education and Technology (ICOMSET 2015)*, 250-255. http://repository.unp.ac.id/616/1/Prosiding%20SR002.pdf

Lee, C., & Ward-Penny, R. (2022). Agency and fidelity in primary teachers' efforts to develop mathematical resilience. *Teacher Development*, *26*(1), 75–93. https://doi.org/10.1080/13664530.2021.2006768

Ling, M. K. D., & Loh, S. C. (2023). Relationships between cognitive pattern recognition and specific mathematical domains in mathematics education. *International Journal of Mathematical Education in Science and Technology*, *54*(2), 159–179. https://doi.org/10.1080/0020739X.2021.1949059

Makarova, O., Ldokova, G., & Egorova, R. (2021). Analysis of students' views of the quality of pedagogical education in Russia. *International Journal of Education in Mathematics, Science and Technology*, *9*(3), 462–481. https://doi.org/10.46328/IJEMST.1624

Malepa-Qhobela, M., & Mosimege, M. (2022). A framework to assist mathematics teachers in integrating problem solving in secondary school classrooms. *Issues in Educational Research*, *32*(4), 1486–1508. https://doi.org/10.3316/informit.806857886364491

Net, W. W. W. P. (2023). The quality of mathematical proficiency in solving geometry problem: difference cognitive independence and motivation. *Pegem Journal of Education and Instruction*, *13*(3), 255–266. https://doi.org/10.47750/pegegog.13.03.27

Öztürk, M., Akkan, Y., & Kaplan, A. (2020). Reading comprehension, mathematics self-efficacy perception, and mathematics attitude as correlates of students' non-routine mathematics problem-solving skills in Turkey. *International Journal of Mathematical Education in Science and Technology*, *51*(7), 1042–1058. https://doi.org/10.1080/0020739X.2019.1648893

Pambudi, D. S., Budayasa, I. K., & Lukito, A. (2020). The role of mathematical connections in mathematical problem solving. *Jurnal Pendidikan Matematika*, *14*(2), 129–144. https://doi.org/10.22342/jpm.14.2.10985.129-144

Pastore, S. (2023). Italian teachers' conceptions of large-scale assessment: what are the implications for assessment literacy? *Teacher Development*, *27*(2), 153–171. https://doi.org/10.1080/13664530.2023.2177719

Plessis, J. Du, & Ewing, B. (2017). Reasonable adjustments in learning programs: Teaching length, mass and capacity to students with intellectual disability. *Universal Journal of Educational Research*, *5*(10), 1795–1805. https://doi.org/10.13189/ujer.2017.051018

Plomp, T. (2013). Educational design research: An introduction. In Plomp, T., & Nieveen, N. (Eds.), *Educational design research - Part A: An introduction,* pp. 10-51. Netherlands Institute for Curriculum Development (SLO). https://slo.nl/publish/pages/2904/educational-design-research-part-a.pdf

Polya, G. (1973). *How to solve it: A new aspect of mathematical method* (2nd ed.). Princeton University Press.

Pongsakdi, N. (2020). What makes mathematical word problem solving challenging? Exploring the roles of word problem characteristics, text comprehension, and arithmetic skills. *ZDM - Mathematics Education*, *52*(1), 33–44. https://doi.org/10.1007/s11858-019-01118-9

Purnomo, E. A., Sukestiyarno, Y. L., Junaedi, I., & Agoestanto, A. (2022). Analysis of problem solving process on HOTS test for integral calculus. *Mathematics Teaching-Research Journal*, *14*(1), 199–214. https://mtrj.commons.gc.cuny.edu/wp-content/blogs.dir/34462/files/2024/07/10Purnomo.pdf

Ramdani, R., Syamsuddin, A., & Sirajuddin, S. (2019). Development of mathematical module-problem solving approach to train student's reflective thinking. *Pedagogical Research*, *4*(4), em0040. https://doi.org/10.29333/pr/5861

Rao, N. J., & Banerjee, S. (2023). Classroom assessment in higher education. *Higher Education for the Future*, *10*(1), 11–30. https://doi.org/10.1177/23476311221143231

Retnawati, H. (2014). *Membuktikan validitas instrumen penelitian.* 31. https://staffnew.uny.ac.id/upload/132255129/pengabdian/7%20Validitas%20Instrumen-alhamdulllah.pdf

Retnawati, H. (2016). *Analisis kuantitatif instrumen penelitian (panduan peneliti, mahasiswa, dan psikometrian)*. Parama publishing.

Retnawati, H., Djidu, H., Kartianom, Apino, E., & Anazifa, R. D. (2018). Teachers' knowledge about higher-order thinking skills and its learning strategy. *Problems of Education in the 21st Century*, *76*(2), 215–230. https://www.scientiasocialis.lt/pec/node/files/pdf/vol76/215-230.Retnawati_Vol.76-2_PEC.pdf

Risma, M., & Yulkifli. (2021). Preliminary study of development of physics e-module using smartphone-assisted inquiry based learning models to support 21stcentury learning. *Journal of Physics: Conference Series*, *1876*(1), 012044. https://doi.org/10.1088/1742-6596/1876/1/012044

Ryoo, J. H., Park, S., Suh, H., Choi, J., & Kwon, J. (2022). Development of a new measure of cognitive ability using automatic item generation and its psychometric properties. *SAGE Open*, *12*(2), 215824402210950. https://doi.org/10.1177/21582440221095016

Saadati, F., & Celis, S. (2022). Student motivation in learning mathematics in technical and vocational higher education: development of an instrument. *International Journal of Education in Mathematics, Science and Technology*, *11*(1), 156–178. https://doi.org/10.46328/ijemst.2194

Saadati, F., Chandia, E., Cerda, G., & Felmer, P. (2023). Self-efficacy, practices, and their relationships; the impact of a professional development program for mathematics teachers. *Journal of Mathematics Teacher Education*, *26*(1), 103–124. https://doi.org/10.1007/s10857-021-09523-2

Schöbel, S. M., Janson, A., & Leimeister, J. M. (2023). Gamifying online training in management education to support emotional engagement and problem-solving skills. *Journal of Management*

Education, *47*(2), 166–203. https://doi.org/10.1177/10525629221123287

Seepiwsiw, K., & Seehamongkon, Y. (2023). The development of mathematical problem-solving and reasoning abilities of sixth graders by organizing learning activities using open approach. *Journal of Education and Learning*, *12*(4), 42-49. https://doi.org/10.5539/jel.v12n4p42

Spinney, J. E. L., & Kerr, S. E. (2023). Students' perceptions of choice-based assessment. *Journal of the Scholarship of Teaching and Learning*, *23*(1), 46–58. https://doi.org/10.14434/josotl.v23i1.31471

Sutiarso, S., Rosidin, U., & Sulistiawan, A. (2022). Developing assessment instrument using polytomous response in mathematics. *European Journal of Educational Research*, *11*(3), 1441–1462. https://doi.org/10.12973/eu-jer.11.3.1441

Tahili, M. H., Tolla, I., Ahmad, M. A., Saman, A., & Samad, S. (2021). The effect of strategic collaboration approach on the national educational standards achievement and service quality in basic education at local government in Indonesia. *Research in Social Sciences and Technology*, *6*(1), 53–82. https://doi.org/10.46303/ressat.2021.4

Tao, Y., Wu, F., Zhang, J., & Yang, X. (2023). Constructing a classroom observation instrument of creative potential for primary school students. *Thinking Skills and Creativity*, *49*, 101317. https://doi.org/10.1016/j.tsc.2023.101317

Turing, A. M. (1954). *Solvable and unsolvable problems*. Penguin Book.

Ulitzsch, E., He, Q., & Pohl, S. (2022). Using sequence mining techniques for understanding incorrect behavioral patterns on interactive tasks. *Journal of Educational and Behavioral Statistics*, *47*(1), 3–35. https://doi.org/10.3102/10769986211010467

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM*, *52*(1), 1–16. https://doi.org/10.1007/s11858-020-01130-4

Vygotsky, L. S. (1994). The problem of the environment. In R. van der Veer, & J. Valsiner (Eds.), *The Vygotsky reader* (pp. 338-354). Black.

Wihardjo, R. S. D., Muktiono, A., Ramadhan, S., Salman, I., & Hadiana, D. (2023). The analysis of instrument quality to measure graduate students' higher order thinking skill in environmental education learning. *International Journal of Membrane Science and Technology*, *10*(3), 277–283. https://doi.org/10.15379/ijmst.v10i3.1528

Yigletu, A., Michael, K., & Atnafu, M. (2023). Professional development on assessment for learning and its effect on pre-service teacher's self-regulated learning. *Cogent Education*, *10*(1), 2222875. https://doi.org/10.1080/2331186X.2023.2222875

Zeighami, M., Malakoutikhah, A., Shahrbabaki, P. M., Al-Oliamat, K., & Dehghan, M. (2024). Development and validation of the nurses sexual harassment scale in Iran. *BMC Nursing*, *23*(1), 107. https://doi.org/10.1186/s12912-024-01759-6

ISSN: 2460-6995 (Online)