❒ 1038

Designing mathematics problem-solving assessment with GeoGebra Classroom: proving the instrument validity

Abdul Haris Rosyidi, Yurizka Melia Sari, Dini Kinati Fardah, Masriyah Masriyah
Department of Mathematics Education, Faculty of Mathematics and Science, Universitas Negeri Surabaya, Surabaya, Indonesia

Article Info

ABSTRACT

Mathematics education is looking for innovative methods to foster problem-solving skills in students. This research develops a problem-solving assessment using GeoGebra Classroom, a versatile interactive mathematics software, to revolutionize mathematics formative assessment and improve students' problem-solving skills. This study adopted the analysis, design, development, implementation, and evaluation (ADDIE) instructional design model stages. The design stage created a comprehensive assessment blueprint, incorporating GeoGebra Classroom functions to create interactive problem-solving tasks. Data analysis used both quantitative and qualitative approaches. Qualitative data consisted of feedback and suggestions from assessment experts, mathematicians, and GeoGebra specialists. Meanwhile, quantitative data included expert scores and cognitive tests that measured students' problem-solving abilities. A cognitive post-test was conducted to measure the progress of students' understanding while using the assessment product. The results of the content validity analysis, assessed using Aiken's V, ranged from 0.85 to 0.92, indicating a high level of validity for the problem-solving skills assessment in terms of content and construction. Some revisions were made to the design of the developed media to make it more interactive for students. These findings suggest that we can further use problem-solving questions integrated with GeoGebra Classroom to uncover the problem-solving skills of junior high school students.

Corresponding Author:

Yurizka Melia Sari
Department of Mathematics Education, Faculty of Mathematics and Science
Universitas Negeri Surabaya
St. Ketintang, Surabaya, Jawa Timur, Indonesia
Email: yurizkasari@unesa.ac.id

1. INTRODUCTION

Mathematics education has always been a subject of continuous exploration and innovation, aiming to enhance students' understanding and application of mathematical concepts in real-world scenarios [1], [2]. It is important to note that effective mathematics education does not only teach formulas and facts but also builds a deep understanding and application of concepts in the context of everyday life. Problem-solving is one of the critical aspects of mathematics learning because it cultivates analytical thinking, logical reasoning, and creativity, all of which are essential skills for students to succeed in mathematics and other aspects of life [2]. Therefore, designing practical assessments to evaluate students' problem-solving skills is paramount in mathematics education.

Mathematical problem-solving tests are an urgent and essential need in mathematics education research instruments and mathematics learning outcomes instruments. Problem-solving tests are needed because it is one of the efforts to improve students' mathematical problem-solving ability as the primary goal

of learning mathematics [3], [4]. In addition, the mathematics problem-solving test is one of the essential components of the independent curriculum [5], [6]. The independent curriculum emphasizes problem-based learning, focusing on students' ability to solve problems effectively and creatively [5]–[7].

In mathematics learning, mathematical problem-solving tests can help teachers evaluate students' ability to understand and apply mathematical concepts in real-life situations and non-routine mathematical problems [3], [4]. Problem-solving tests can also help students develop critical, creative, and innovative thinking skills for solving complex problems [8]. In addition, math problem-solving tests can also help teachers evaluate the effectiveness of learning and the success of implementing an independent curriculum [9]. With problem-solving tests, teachers can determine the extent of students's ability to solve problems related to mathematical concepts learned for follow-up activities.

In recent years, technology has become a powerful tool to revolutionize how educators teach and students learn. One such technological advancement is GeoGebra Classroom, an interactive mathematics software that integrates various mathematical representations, including algebra, geometry, calculus, and statistics, to foster a dynamic learning environment [10]–[12]. Using GeoGebra Classroom enables teachers to create interactive lessons and assessments that actively engage students in learning, promoting a more profound understanding and meaningful application of mathematical concepts [11].

There are several technologies focused on math tests that can be used to represent, explore, and perform math tasks: dynamic geometry software (DGS), computational and representation tools (Wolfram's alpha) [13], Microworlds, or computer simulations [14]), and assistive technologies that are useful for explaining, sharing, and discussing mathematical ideas or problems (communication applications such as Skype or FaceTime and presentation technologies such as Keynote or PowerPoint). There are also online platforms that include videos to explain mathematical themes, examples of proposed problems and tasks (https://www.khanacademy.org/coach/dashboard), or online development (https://www.wikipedia.org) where students can consult information about content, concepts, or events [15], [16]. All these technological developments provide different conveniences for teachers and students to work on math problems. The aim is for students to use them throughout all their learning experiences.

Several studies have shown positive results regarding integrating technology in math problem-solving tests. It revealed that the problem-solving ability of high school students who were given GeoGebra integration with a problem-based approach was better than that of those who were given conventional tests [17]. In addition, integrating assessment with technology offers opportunities to design rich, dynamic, and interactive items [18]. The digital assessment facilitates items that allow students to manipulate objects, explore non-fixed properties, or create examples. Additionally, we can design items with multiple correct answers or completion strategies, and use various item types. Appropriate feedback designs and scoring rules or grading schemes can be created and implemented [19].

This article focuses on developing a novel assessment design that integrates GeoGebra Classroom to evaluate students' problem-solving skills in mathematics. The objective is to harness the potential of technology to enhance traditional assessment methods and foster a more comprehensive and immersive learning experience for students. Consequently, it is necessary to research the design of a prototype of the DGS-integrated mathematics problem-solving skills assessment, namely the GeoGebra Classroom. This research is expected to present an alternative technology-integrated test focusing on problem-solving. The specific objectives of the research to be carried out are to produce validity for the problem-solving test of the DGS integrated problem-solving test.

2.   METHOD
2.1. Research design
This study's research methodology employs the analysis, design, development, implementation, and evaluation (ADDIE) development model, which comprises five sequential stages: analysis, design, development, implementation, and evaluation [20]. The ADDIE model provides a systematic framework for designing and evaluating teaching materials or learning media. The analysis stage is used to understand learning needs and objectives; design includes detailed planning; development involves making materials; implementation is the stage of implementation in the field; and evaluation focuses on assessing the effectiveness and improvement of learning content. This study's application of the ADDIE model ensures an in-depth analysis-based teaching material development process and continuous evaluation of implementation results. The study visualizes the five stages, as shown in Figure 1.

The analysis phase involves exploring learning objectives and understanding the educational context. The design phase further focuses on creating a comprehensive assessment blueprint, where integrating GeoGebra Classroom features plays an important role in designing interactive problem-solving tasks that challenge students to apply mathematical concepts in real-world scenarios.
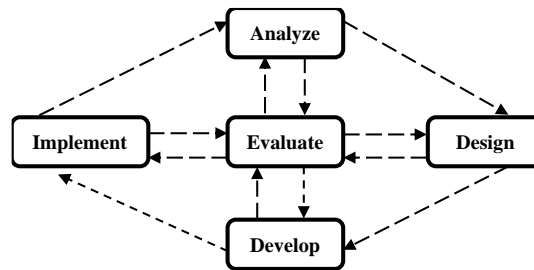
Figure 1. ADDIE development model

At the development stage, this research aims to prove the validation of the assessment. The draft instrument has been validated to obtain expert judgment. The instrument that has proven its validity will be tested on a limited basis with some junior high school students in Surabaya, East Java. The product trial design in this study involves a content validation step to prove the validity of each item developed. This step is carried out to ensure the accuracy of the items developed.

2.2. Content validity

In this study, data were obtained from the results of instrument validation, expert judgment, and student responses to the readability of assigned math assignments or problems. The researchers processed the data from the instrument validation results quantitatively, taking into account the expert comments, and then made improvements to the instrument until they declared it valid and feasible for use. The data from the results of the readability of the instrument by students was processed qualitatively and using certain criteria so that the instrument was declared to need improvement or not.

Expert consensus determines the validity. Experts in the relevant field assess the domain to determine the content's validity [21], [22]. Experts deem tools like tests or questionnaires valid when they agree that they accurately measure the mastery of a specified domain. We use Aiken's formula to determine accuracy. Aiken developed the formula V to compute the content-validity coefficient, grounded on evaluations from an expert panel of n individuals regarding how accurately an item reflects the targeted domain. Aiken illustrates his formula (1) [22] as follows:

$$v = \frac{\sum s}{n(c-1)} \tag{1}$$

where,

v   =   Aiken's index validity
s   =   r-lo
lo   =   lowest validity score
c   =   highest validity score
r   =   score given by rater
n   =   the number of rater

One can assess the validity of an item or instrument based on three categories by calculating the V index. If the V index is less than or equal to 0.4, validity is considered low, indicating limitations in measuring the intended concept. If the index ranges from 0.4 to 0.8, the validity is considered moderate, indicating an adequate level of measurement conformity. Meanwhile, suppose the V index is more significant than 0.8. In that case, the validity is considered very high, indicating a strong ability of the item or instrument to measure the concept accurately and consistently [22].

3. RESULTS AND DISCUSSION

3.1. The problem-solving instruments

This instrument, focusing on junior high school students, aims to test and uncover their mathematical problem-solving skills using GeoGebra Classroom. Additionally, the instrument's design aims to gauge students' technological proficiency. The instrument also aims to support the development of students' problem-solving skills. Furthermore, it helped identify specific areas that required further attention, allowing for a more targeted approach to learning. The instrument comprises 7 questions, each designed to align with the learning outcomes of the independent curriculum in phase D and tailored to various problem-solving scenarios. Table 1 shows the results of the problem grids that were developed.

Table 1. The framework of mathematics problem solving integrated with GeoGebra Classroom

Learning Outcomes	Topic	Problems Type	Designed Tasks
Numbers By the end of phase D, learners can read, write and compare whole numbers, rational and irrational numbers, decimal numbers, whole numbers and roots, and numbers in scientific notation. They can apply arithmetic operations to real numbers and provide estimates in problem-solving (including concerning financial literacy). In problem-solving, learners can use prime factorization and ratio (scale, proportion, and rate of change).	Integers	Finding and proving	Find properties of integers that can be expressed as the sum of consecutive integers
Algebra By the end of phase D, learners can recognize, predict and generalize patterns in the form of arrangements of objects and numbers. They can express a situation in algebraic form. They can use the properties of operations (commutative, associative, and distributive) to produce equivalent algebraic forms. Learners can understand relations and functions (domain, codomain, range) and present them as arrow diagrams, tables, sets of ordered pairs, and graphs. They can distinguish some nonlinear functions from linear functions graphically. They can solve linear equations and inequalities of one variable. They can present, analyze and solve problems using relations, functions, and linear equations. They can solve systems of linear equations of two variables through several ways of problem-solving.	Linear Function	Finding	Find the characteristics of the graph of the function $f(x)=mx+c$ with the following categories (1) $m>0$ and $c>0$ (2) $m>0$ and $c<0$ (3) $m>0$ and $c=0$ (4) $m<0$ and $c>0$ (5) $m<0$ and $c<0$ (6) $m<0$ and $c=0$ by pairing in each category (some pairs are true and some are false) with arguments
		Open-Ended	Find the general form of the linear function $f(x)=mx+c$ through the point $(1,2)$.
Measurement By the end of phase D, learners can explain how to determine the area of a circle and solve related problems. They can explain how to determine the surface area and volume of spatial shapes (prisms, tubes, balls, pyramids, and cones) and solve related problems. They can explain the effect of proportional changes of flat and spatial shapes on the length, angle magnitude, area, and volume.	Area of an equilateral triangle	Proving	Based on the exploration above, Abdul thinks that to get a triangle with an area of 2 times the original triangle, then what can be done is to double the length of each side of the triangle. Is this opinion right or wrong? Please share your opinion on this matter.
	Rectangle	Proving	The area and perimeter of a rectangle will remain the same when the length is doubled, and the width is halved. Is it always true, sometimes accurate, or not true? Explain your reasoning.
Geometry By the end of phase D, learners can construct the nets of spatial figures (prisms, cubes, pyramids, and cones) and build the figures from their nets. Learners can use the relationship between angles formed by two intersecting lines and by two parallel lines cut by a transversal line to solve problems (including determining the sum of the angles in a triangle and the magnitude of an unknown angle in a triangle). They can explain the properties of congruence and similarity in triangles and quadrilaterals and use them to solve problems. They can demonstrate the correctness of the Pythagorean theorem and use it in solving problems (including the distance between two points on the Cartesian coordinate plane). Learners can perform a single transformation (reflection, translation, rotation, and dilation) of points, lines, and flat figures on the Cartesian coordinate plane and use it to solve problems.	Reflection on the line $x=h$	Finding	Through exploring the reflection of a point on the line $x=h$ in GeoGebra, students can determine some of the following problems: Determine the image of $A(21,22)$ with respect to $x=10$ Determine h if point (a,b) becomes point (a', b') Determine the shadow of $A(a,b)$ with respect to $x=h$
Statistics and probability By the end of phase D, learners can formulate a question and collect, present, and analyze data to answer the question. They can use bar charts and pie charts to present and interpret data. They can take a representative population sample to obtain data about them and their environment. They can determine and interpret the mean, median, mode, and range of data to solve problems (including comparing data to its group, comparing two data groups, predicting, and making decisions). They can investigate possible changes to the center measurement due to changes in the data. Learners can explain and use the notions of probability and relative frequency to determine the expected frequency of an event in a simple experiment (all outcomes can occur equally).	Mean	Proving	Prove that the average of a set of numbers will increase by n if each number in the set is increased by n?

Here is one example of a problem that measures maths problem-solving skills integrated with the GeoGebra Classroom problem-solving skills, as shown in Figure 2. In this context, students are allowed to conduct an interactive exploration of the characteristics of the graph of the linear function $f(x)=mx+c$, where

m is the slope of the line and c is the intersection with the y-axis. By changing the values of m and c, students can immediately see the impact of these changes on the shape and position of the graph. For example, when changing m, they will observe how the line becomes steeper or flatter depending on whether the value of m is positive or negative. Similarly, when changing c, they will observe how the position of the line changes concerning the y-axis.

This exploration can help students better understand the relationship between the parameters m and c and the shape of the graph [23], [24]. They can relate mathematical concepts such as slope and the intersection of axes y-axis with the concrete visualizations they produce. Through this interactive experience, students can internalize these concepts better than simply listening to theoretical explanations. In addition, this exploration can also arouse students' curiosity and curiosity toward the mathematical relationships underlying the graphs of linear functions [25]. This way, using tools such as GeoGebra can help create more engaging, effective, and interactive learning in understanding basic mathematical concepts such as linear functions [25].

After students undertake an interactive exploration by manipulating the values of m and c in the equation f(x)=mx+c, they are confronted with questions that encourage the development of problem-solving skills and deep understanding, as shown in Figure 3. One such question is how to describe the characteristics of the f(x)=mx+c graph when m>0 and c>0. In answering this question, students must apply their understanding of slope (m) concepts and intersection with the y-axis (c). When m>0, they will realize that the graph will have a positive slope, indicating that the line will tend to rise from the bottom left to the top right. It indicates that the larger the value of x, the larger the value of f(x), which corresponds to a steady and steady increase in this relationship. On the other hand, when c>0, students will notice that the graph intersects the y-axis at a favorable point on the y scale, indicating that the line has a positive y-value when x=0.

Students can develop analytical and communicative thinking skills by responding to such questions. They can articulate their understanding of the relationship between the values of m and c with a graph and exercise their ability to think logically and formulate arguments. In addition, this kind of questioning also helps students develop their visual analytical acuity and ability to relate mathematical concepts to the graphical representations they generate through GeoGebra [26].

3.2. Proving the content validity

At the validation stage of the problem-solving skills test instrument, the researcher evaluates the test that has been developed. The test was assessed based on several aspects, namely material aspects, construction aspects, language aspects, and media aspects. Each test item is given a score based on these four aspects of assessment. The assessment was conducted by two mathematics education assessment experts and one mathematics expert, who answered the questions and gave scores based on predetermined assessment aspects, as presented in Table 2.
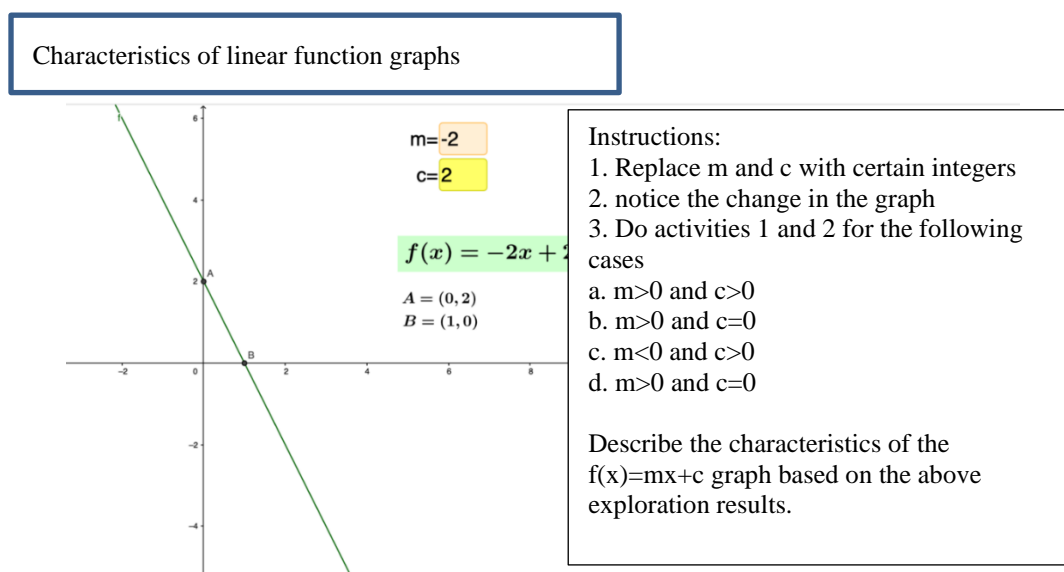

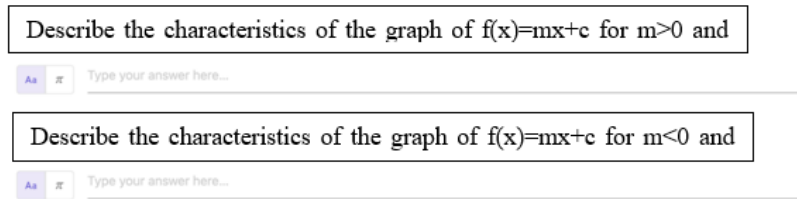Figure 2. An example of a problem integrated with GeoGebra Classroom linear function topic

Describe the characteristics of the graph of f(x)=mx+c for m>0 and

Aa π   Type your answer here...

Describe the characteristics of the graph of f(x)=mx+c for m<0 and

Aa π   Type your answer here...

Figure 3. Some sample problems were provided to students after their exploration

Table 2. The Results of content validity

Items	Expert 1	Expert 2	Expert 3	Validity index
1	4	3	4	0.89
2	4	3	3	0.78
3	3	3	3	0.67
4	3	3	4	0.78
5	4	3	4	0.89
6	4	3	4	0.89
7	4	3	4	0.89

This process's results indicate the validity level of the developed problem-solving skills test. Expert views express this validity score in the form of Aiken's V index, measuring the extent to which the test items can be considered valid. The Aiken's V validity score range is 0.37 to 1.00, and in this study, the Aiken's V index value for seven items ranged from 0.67 to 0.89. According to the assessment that has been conducted, all items are considered valid and suitable for use in the next stage of the empirical trial [21], [22], and [27].

This result is in line with the interpretation steps used by other researchers [28], where if the item validity score is within the range of 0.37 to 1.00, then the items are considered valid. Thus, this finding provides confidence that the test instrument developed through this validation stage has passed a careful evaluation process and is ready to be empirically tested to measure problem-solving skills in the intended subjects.

3.3. Proving validity based on the student response process

Validity based on student response processes focuses on understanding how students think and respond during the assessment process. In an educational context, this relates to how students' approach, process, and answer questions in a test or assessment instrument. Researchers analyse students' responses to each item or question to ensure an assessment instrument is valid in this context.

GeoGebra Classroom allows for the recording of students' problem-solving processes, as shown in Figure 4. Students have tried to replace the numbers on the left and observe the average change after adding a certain number on the right. However, the initial draft of GeoGebra Classroom needs to embed the questions students answer. So that students write their answers on paper, as shown in Figure 5.
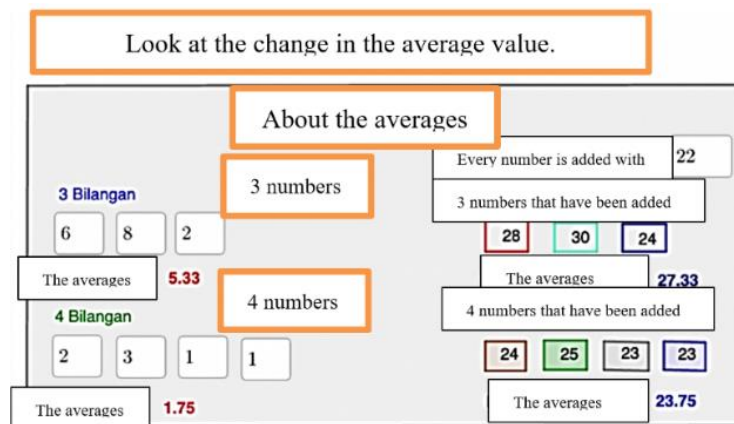

Figure 4. Results of student work in GeoGebra Classroom

In addition, the student also revealed that he needed clarification about what was asked from this problem because no questions directed students to solve the average topic problem. It can be seen from the student's answer, which is still general and does not lead to the specific characteristics of the average of a number when summed with several ns. Therefore, this question needs to be revised. Therefore, this question needs to be revised by adding questions that can direct students to conclude according to the learning objectives. The revision is made by adding questions after students explore GeoGebra, as shown in Figure 6.

Translate: b1 is equal to the front result of the average. If all numbers are the same, the average is the same. If b2 and b3 are added, the average is also increased. If one of the numbers is increased, the result is increased.

Figure 5. Student's conclusion on paper

Figure 6. The results of the revision of the question according to the validation evidence of the student work process

It ensures that students can solve the problems presented in the GeoGebra Classroom. By understanding how students process information and make decisions when answering problems, educators can ensure that the problem measures the intended construct and that no other factors (e.g., reading ability or guessing strategies) are present [29]. Therefore, relying on response processes for validity enhances the assurance that the assessment tool accurately captures the ability or knowledge under evaluation, instead of focusing on irrelevant variables [30], [31].

4. CONCLUSION

The final product of this research is a valid assessment instrument for problem-solving skills, integrated with the GeoGebra Classroom, based on the results and discussions. The development process involved the definition, design, and development of this instrument. The assessment instrument consists of seven essay questions that have been declared to have appropriate content validity. The instrument developed has an Aiken validity coefficient of 0.67–0.89 (high). Therefore, the developed instrument can be used as an accurate learning assessment tool or assessment of problem-solving skills. Some suggestions that can be conveyed based on the results of this study in the context of developing learning assessment instruments include the following: In junior high school mathematics studies, teachers can use the problem-solving skills assessment instrument as an evaluation tool. Teachers can also develop mathematics problem-solving skills on other topics. Other researchers can use the results of this study as a reference for integrating problem-solving-based mathematics learning with technology.

ACKNOWLEDGEMENTS

REFERENCES

[1] V. Koshy, P. Ernest, and R. Casey, "Mathematically gifted and talented learners: theory and practice," *International Journal of Mathematical Education in Science and Technology*, vol. 40, no. 2, pp. 213–228, Mar. 2009, doi: 10.1080/00207390802566907.

[2] P. Yuanita, H. Zulnaidi, and E. Zakaria, "The effectiveness of Realistic Mathematics Education approach: The role of mathematical representation as mediator between mathematical belief and problem solving," *PLOS ONE*, vol. 13, no. 9, Sep. 2018, doi: 10.1371/journal.pone.0204847.

[3] P. Pardimin and S. A. Widodo, "Increasing skills of student in junior high school to problem solving in geometry with guided," *Journal of Education and Learning (EduLearn)*, vol. 10, no. 4, pp. 390–395, Nov. 2016, doi: 10.11591/edulearn.v10i4.3929.

[4] I. W. Widiana, I. G. N. Japa, I. M. Suarjana, and K. Sujendra Diputra, "the students' ability to solve realistic mathematical problems through polya type problem solving learning model," *Journal of Education and Learning (EduLearn)*, vol. 12, no. 3, pp. 399–405, Aug. 2018, doi: 10.11591/edulearn.v12i3.4526.

[5] A. B. Perdana, S. N. Indah, Y. W. Kusumo, and K. A. Ardhanariswari, "How does the diffusion of Kampus Merdeka innovation work at the university?," *Proceedings of the International Conference on Engineering, Technology and Social Science (ICONETOS 2020)*, 2021, pp. 469-476, doi: 10.2991/assehr.k.210421.068.

[6] R. Christanti and A. A. Sukoco, "Freedom to learn–independent campus policy: do we really find our freedom?," *Journal of Education and Learning (EduLearn)*, vol. 16, no. 2, pp. 189–198, May 2022, doi: 10.11591/edulearn.v16i2.20477.

[7] E. T. Alawiah and D. Setyorini, "Use of The Merdeka teaching platform application in the implementation of the independent curriculum," *International Journal of Information System & Technology Akreditasi*, vol. 7, no. 3, pp. 162–167, 2023.

[8] T. M. Sholihah and B. Lastariwati, "Problem based learning to increase competence of critical thinking and problem solving," *Journal of Education and Learning (EduLearn)*, vol. 14, no. 1, pp. 148–154, Feb. 2020, doi: 10.11591/edulearn.v14i1.13772.

[9] H. Hobri, I. K. Ummah, N. Yuliati, and D. Dafik, "The effect of jumping task based on creative problem solving on students' problem solving ability," *International Journal of Instruction*, vol. 13, no. 1, pp. 387–406, Jan. 2020, doi: 10.29333/iji.2020.13126a.

[10] M. N. Isman, "Utilization of the Geogebra program in mathematics learning, in Indonesia: Pemanfaatan Program Geogebra Dalam Pembelajaran Matematika," *Delta-pi : Jurnal Matematika Dan Pendidikan Matematika*, vol. 5, no. 1, pp. 10–19, 2016.

[11] J. Zöchbauer, M. Hohenwarter, and Z. Lavicza, "Evaluating GeoGebra Classroom with usability and user experience methods for further development," *International Journal for Technology in Mathematics Education*, vol. 28, no. 3, pp. 183–191, 2021.

[12] I. Benning, C. Linsell, and N. Ingram, "Examining the changes in mathematics teachers' technology dispositions through GeoGebra-mediated professional development," *Asian Journal for Mathematics Education*, vol. 2, no. 1, pp. 42–63, 2023, doi: 10.1177/27527263231163276.

[13] S. Abramovich, "Using Wolfram alpha with elementary teacher candidates: from more than one correct answer to more than one correct solution," *Mathematics*, vol. 9, no. 17, p. 2112, Sep. 2021, doi: 10.3390/math9172112.

[14] H. M. Hodgetts, S. Packwood, F. Vachon, and S. Tremblay, "A microworld simulation of dynamic cognition as a test of executive function," *Journal of Clinical and Experimental Neuropsychology*, vol. 45, no. 2, pp. 165–181, Feb. 2023, doi: 10.1080/13803395.2023.2214297.

[15] H. T. Hou, K. E. Chang, and Y. T. Sung, "Analysis of problem-solving-based online asynchronous discussion pattern," *Educational Technology and Society*, vol. 11, no. 1, pp. 17–28, 2008.

[16] C. Battistella and F. Nonino, "Open innovation web-based platforms: The impact of different forms of motivation on collaboration," *Innovation*, vol. 14, no. 4, pp. 557–575, Dec. 2012, doi: 10.5172/impp.2012.14.4.557.

[17] A. Septian and E. Monariska, "The improvement of mathematics understanding ability on system of linear equation materials and students learning motivation using geogebra-based educational games," *Al-Jabar : Jurnal Pendidikan Matematika*, vol. 12, no. 2, pp. 371–384, 2021, doi: 10.24042/ajpm.v12i2.9927.

[18] J. W. Pellegrino and E. S. Quellmalz, "Perspectives on the integration of technology and assessment," *Journal of Research on Technology in Education*, vol. 43, no. 2, pp. 119–134, Dec. 2010, doi: 10.1080/15391523.2010.10782565.

[19] Y. M. Sari, B. Kartowagiran, H. Retnawati, and S. Fiangga, "The characteristics of digital assessment bloom for indonesian junior high school," *Proceedings of the Mathematics, Informatics, Science, and Education International Conference (MISEIC 2018)*, 2018, pp. 26-30, doi: 10.2991/miseic-18.2018.7.

[20] J. M. Spector, T. E. Johnson, and P. A. Young, "An editorial on research and development in and with educational technology," *Educational Technology Research and Development*, vol. 62, no. 1, pp. 1–12, Feb. 2014, doi: 10.1007/s11423-014-9331-z.

[21] L. R. Aiken, "Three coefficients for analyzing the reliability and validity of ratings," *Educational and Psychological Measurement*, vol. 45, no. 1, pp. 131–142, Mar. 1985, doi: 10.1177/0013164485451012.

[22] H. Retnawati, "Proving content validity of self-regulated learning scale (the comparison of Aiken index and expanded Gregory index)," *Research and Evaluation in Education*, vol. 2, no. 2, pp. 155–164, Dec. 2016, doi: 10.21831/reid.v2i2.11029.

[23] P. Sari, A. Hadiyan, and D. Antari, "Exploring derivatives by means of GeoGebra," *International Journal on Emerging Mathematics Education*, vol. 2, no. 1, pp. 65–78, Feb. 2018, doi: 10.12928/ijeme.v2i1.8670.

[24] I. M. Romero Albaladejo and M. del M. García López, "Mathematical attitudes transformation when introducing GeoGebra in the secondary classroom," *Education and Information Technologies*, pp. 1–26, 2023, doi: 10.1007/s10639-023-12085-w.

[25] E. Nzaramyimana, "Effectiveness of GeoGebra towards students' active learning, performance and interest to learn mathematics," *International Journal of Mathematics and Computer Research*, vol. 09, no. 10, pp. 2423–2430, 2021, doi: 10.47191/ijmcr/v9i10.05.

[26] D. Abrahamson and R. Abdu, "Towards an ecological-dynamics design framework for embodied-interaction conceptual learning: the case of dynamic mathematics environments," *Educational Technology Research and Development*, vol. 69, no. 4, pp. 1889–1923, Aug. 2021, doi: 10.1007/s11423-020-09805-1.

[27] H. Narendrany, B. Hayat, and W. Rahayu, "Assessment of the validity and reliability of mental health instruments of high school student in Indonesia," *European Journal of Educational Research*, vol. 10, no. 2, pp. 729–742, Apr. 2021.

[28] D. N. Kowsalya, V. Lakshmi, and K. P. Suresh, "Development and validation of a scale to assess self-concept in mild intellectually disabled children," *International Journal of Social Sciences and Education*, vol. 2, no. 4, pp. 2223–4934, 2012.

[29] T. Suryani, W. Rahayu, and A. Saptono, "Development and validation technological pedagogical content knowledge (tpack) instrument for teacher mathematics in elementary school," *International Journal of Multicultural and Multireligious Understanding*, vol. 8, no. 8, pp. 445–457, Aug. 2021, doi: 10.18415/ijmmu.v8i8.2951.
[30] S. Shankar, C. St-Onge, and M. E. Young, "When I say … response process validity evidence," *Medical Education*, vol. 56, no. 9, pp. 878–880, Sep. 2022, doi: 10.1111/medu.14853.
[31] M. S. B. Yusoff, "ABC of content validation and content validity index calculation," *Education in Medicine Journal*, vol. 11, no. 2, pp. 49–54, Jun. 2019, doi: 10.21315/eimj2019.11.2.6.

BIOGRAPHIES OF AUTHORS

Abdul Haris Rosyidi [ID] [G] [SC] [◉] is a lecturer in mathematics education at Universitas Negeri Surabaya. He completed his bachelor's and master's degrees in mathematics education at Universitas Negeri Surabaya. Currently, he is a doctoral candidate in mathematics education and is pursuing his doctoral studies at Universitas Negeri Malang. Apart from being a lecturer, he is also a book writer and speaker on mathematics education. His research areas are problem-solving, critical thinking and evaluation. He can be contacted at email: abdulharis@unesa.ac.id.

Yurizka Melia Sari [ID] [G] [SC] [◉] is a lecturer at department of mathematics education, Universitas Negeri Surabaya (UNESA). She obtained bachelor's degree in department of mathematics education, Universitas Negeri Surabaya in 2012 and finished her master's degree program in mathematics education at Universitas Negeri Surabaya in 2014. She pursued doctoral degree in educational research and evaluation in 2020 from Universitas Negeri Yogyakarta. Her area of interest includes mathematics education, evaluation, assessment, reasoning, and classroom research. She can be contacted at email: yurizkasari@unesa.ac.id

Dini Kinati Fardah [ID] [G] [SC] [◉] is a lecturer in mathematics education at Universitas Negeri Surabaya. She completed her Bachelor of Mathematics Education at Universitas Negeri Yogyakarta and her Master of Mathematics Education at Universitas Negeri Semarang. Currently, she is a doctoral candidate in mathematics education and is pursuing her doctoral studies at UPI Bandung. Her research areas are creative thinking, critical thinking, and evaluation. She can be contacted at email: dinikinati@unesa.ac.id

Masriyah Masriyah [ID] [G] [SC] [◉] is a Senior Asssociate Professor at Mathematics education department, Universitas Negeri Surabaya (UNESA). She received bachelor's degree in department of mathematics education, IKIP Surabaya in 1984, received master degree in department of mathematics education, IKIP Malang in 1988, and she obtained doctoral degree in mathematics education in 2014 from Universitas Negeri Surabaya (UNESA). Her research interests lie in the teacher and teacher education, assessment and evaluation, classroom research, and problem posing. She can be contacted at email: masriyah@unesa.ac.id.