**Journal of Mathematics Education**
Volume 15, No. 3, 2026
https://doi.org/10.22460/infinity.v15i3.p717-756

p-ISSN 2089-6867
e–ISSN 2460-9285

# Development of ARGEONET integrated with metaverse and M6 learning model to enhance students' mathematical critical thinking skills

**Puguh Darmawan[1*], Imam Rofiki[1], Tajularipin Sulaiman[2], Ardi Dwi Susandi[3], Thesa Kandaga[3], Siti Fatimah[4], Entit Puspita[5]**

[1]Department of Mathematics, Universitas Negeri Malang, East Java, Indonesia
[2]Department of Foundations of Education, Universiti Putra Malaysia, Serdang, Malaysia
[3]Department of Mathematics Education, Universitas Terbuka, Banten, Indonesia
[4]Department of Mathematics Education, Universitas Pendidikan Indonesia, West Java, Indonesia
[5]Department of Master of Pedagogy, Universitas Nusa Putra, West Java, Indonesia
[*]Correspondence: puguh.darmawan.fmipa@um.ac.id

**Received**: Nov 24, 2025 | **Revised**: Jun 13, 2026 | **Accepted**: Jun 17, 2026 | **Published Online**: Jul 20, 2026

## Abstract

Critical thinking is an essential 21st-century competency. Many junior high school students still demonstrate low levels of mathematical critical thinking. This study aims to develop ARGEONET, an integrated augmented reality media resource that supports the MathVerse M6 learning model and is designed to improve students' critical thinking skills in mathematics. The research employed a Research and Development (Plomp's educational design framework). The participants were 30 junior high school students in Banyuwangi Regency, East Java, Indonesia, selected through purposive sampling. Data were collected using expert validation sheets, observation sheets, student response questionnaires, and mathematical critical-thinking tests. The results show that ARGEONET obtained very high validity, with scores of 89% and 93% from validators, and was classified as very practical, with scores of 86.06% from students and 91.43% from teachers. Effectiveness analysis showed significant differences in five critical thinking indicators: focus (sig. 0.004), reason (sig. 0.001), inference (sig. 0.003), situation (sig. 0.050), clarity (sig. 0.481), and overview (sig. 0.042). The effect size from each indicator, namely small (clarity), medium (situation and overview), and large (focus, reason, inference). ARGEONET shows potential as an augmented reality media to enhance students' critical thinking skills.

**Keywords**:
Augmented reality, Critical thinking, Learning media, Metaverse, Technology

**How to Cite**:
Darmawan, P., Rofiki, I., Sulaiman, T., Susandi, A. D., Kandaga, T., Fatimah, S., & Puspita, E. (2026). Development of ARGEONET integrated with metaverse and M6 learning model to enhance students' mathematical critical thinking skills. *Infinity Journal*, *15*(3), 717-756. https://doi.org/10.22460/infinity.v15i3.p717-756

## 1. INTRODUCTION

Critical thinking skills are among the essential competencies of the 21st century. In mathematics education, individuals with strong critical thinking skills tend to be more proficient in problem-solving (Aslan & Aybek, 2024; Susandi et al., 2022). Critical thinking skills enable students to solve mathematical problems while making logical and rational decisions based on the available evidence and contextual aspect (Demircioglu et al., 2023; Mahdi & Sahari, 2024). Therefore, the Indonesian government emphasizes the importance of developing critical thinking skills in every teaching and learning activity to prepare students to face future challenges (Kemendikbud, 2022). To facilitate the development and evaluation of these competencies, many experts have formulated aspects or indicators of critical thinking that can be used as a basis for planning learning activities. Several experts who formulated aspects of critical thinking are Facione et al. (1995) and Ennis (1996).

Facione et al. (1995) divided critical thinking into six abilities, which include: (1) interpretation, the ability to understand and restate the information obtained; (2) analysis, the ability to observe and break down information in detail for further examination; (3) inference, the ability to draw conclusions; (4) evaluation, the ability to assess the strength of arguments; (5) explanation, explaining the thought process or reasoning used to reach a conclusion; and (6) self-regulation, reflecting on one's thinking and correcting errors if necessary. Meanwhile, Ennis (1996) states that critical thinking consists of six components which is known by the acronym FRISCO, namely (1) focus, concentrating attention on the problem at hand; (2) reason, the ability to provide appropriate and logical justification; (3) inference, which is the ability to draw conclusions based on evidence or information; (4) situation, which is the ability to adjust judgments according to the context of the problem; (5) clarity, or the ability to explain ideas clearly and in detail; and (6) overview, which is the ability to review the thinking process that has been undertaken. The examination of critical thinking skills in this study uses indicators developed by Ennis (1996). This selection is due to the Ennis indicators enabling researchers to identify the stage of critical thinking in students' problem-solving outcomes. These six components illustrate an individual's capacity for rational, directed, and reflective thinking when addressing a problem. Based on the six components of critical thinking according to Ennis (1996), indicators of critical thinking ability are formulated as presented in Table 1.

**Table 1**. Indicators of critical thinking ability

| No | Critical Thinking Ability | Indicators |
|---|---|---|
| 1 | Focus | Writing down the necessary information to answer questions |
| 2 | Reason | Providing reasoning based on relevant facts or evidence at every step of decision-making |
| 3 | Inference | Drawing correct conclusions based on appropriate reasoning |
| 4 | Situation | Using all information according to the given problem |
| 5 | Clarity | Providing further explanation of the conclusions that have been made |
| 6 | Overview | Checking for errors that occur at each step that has been taken |

Individuals who can achieve the six components of critical thinking ability in Table 1 are considered capable of producing comprehensive problem-solving, from focusing attention on the problem to reviewing the thought process that has been carried out. However, in reality, not all individuals, especially students, can reach these indicators. This is due to students' critical thinking skills in mathematics still being relatively low. The low level of mathematical critical thinking ability is shown by the results of the 2015 Trends in International Mathematics and Science Study (TIMSS), which placed Indonesia at rank 46 out of 51 countries (Egan et al., 2022; Whitney-Smith, 2023). The results of the 2022 Programme for International Student Assessment (PISA) also show that Indonesian students' mathematical literacy ranks 69th out of 80 countries (OECD, 2023; Wijaya et al., 2024). In addition, based on a preliminary study of junior high school students in Banyuwangi Regency, East Java, Indonesia, the three indicators of critical thinking in solving problems related to nets of flat-sided three-dimensional shapes have not yet been met. The indicators that have not yet been met are focus, situation, and clarity. First, the focus indicator was not met because the students were unable to write down the information needed to answer the question, namely the length, width, and height of the block. Second, the situation indicator was not met because the students were unable to use all the available information to solve the problem. Third, the clarity indicator was not met because the students were unable to explain the reasoning behind their final results and conclusions. These results highlight the low level of Indonesian students' mathematical critical thinking skills, indicating the need for learning designs that can foster such abilities. Integrating digital technology into mathematics learning is one of the potential solutions to address the low level of students' mathematical critical thinking skills. This integration can be achieved through the use of Augmented Reality (AR). AR has the potential to support the development of students' critical thinking skills (Demircioglu et al., 2023). Interactive visualizations in AR encourage students to analyze information, evaluate alternative solutions, and engage in logical reasoning to solve problems (Mandala et al., 2025; Maulida et al., 2024).

The potential of AR to support students' critical thinking skills is closely linked to the increasing role of digital technology in education. In the digital era, students often described as digital natives are accustomed to technology from an early age (Alruthaya et al., 2021), making technology integration essential to enhance mathematics learning. One of promising technology-based educational innovations is the Metaverse. The Metaverse is a virtual environment that enables interactive, immersive, and collaborative learning experiences (Cheerapakorn et al., 2024; Ma et al., 2023; Talan & Kalinkara, 2022). The integration of the Metaverse into mathematics education can enhance students' motivation and engagement in learning activities (Al Khateeb & Alotaibi, 2024; Kongkhen & Chatwattana, 2023; Suh & Ahn, 2022). Since mathematical concepts are often abstract and difficult for students to grasp, Metaverse-based visualization can help make these concepts more concrete and easier to understand (Avcu & Yaman, 2024; Tural & Koçak, 2023). Thus, interactive virtual environments like the Metaverse not only strengthen students' conceptual understanding but also enhance their mathematical critical thinking (Hazaymeh & Alomery, 2021; Pereles et al., 2024). To optimize the potential of the Metaverse in fostering students' mathematical critical thinking skills, it is necessary to integrate it with a learning model that emphasizes active and collaborative.

The M6 learning model is designed to integrate various syntaxes that involve active and collaborative student activities. This model consists of six main phases: (1) focusing on prior abilities, (2) justifying concepts, (3) investigating problems, (4) presenting ideas, (5) conducting evaluation, and (6) concluding the learning outcomes (Susandi et al., 2022). The purpose of this model is to develop critical thinking, creativity, collaboration, and communication, key competencies required to face 21st-century challenges. By integrating Metaverse technology into the M6 learning model, classroom learning is expected to become more interactive, engaging, and effective, thereby enhancing students' mathematical critical thinking skills. Based on previous studies, research on the integration of Metaverse technology in mathematics education demonstrates significant potential for enhancing learning quality. The position of the present study is summarized in Table 2.

**Table 2**. Position of this research

| Researcher (Year) | Focus | Result |
|---|---|---|
| Avci and AkgÜL (2024); Mutchima et al. (2025); Rachmadtullah et al. (2023) | Development metaverse to enhance students' critical thinking | Metaverse can create an interactive and immersive learning environment. Furthermore, the use of Metaverse in education can significantly improve learning quality, including critical thinking |
| Chang et al. (2022); Eşin and Özdemir (2022) | Development of Augmented Reality (AR) to enhance students' understanding | Augmented reality enriches learning experiences by enhancing student engagement and understanding, particularly in visualizing complex concepts like spatial geometry and function graphs |
| Khusna et al. (2024); Nosrat and Parsafar (2018); Wang and Abdullah (2024) | Students' critical thinking in solving mathematical problems | A well-designed mathematics learning approach that integrates problem-solving methods, collaborative non-routine tasks, and instructional strategies with clearly defined critical thinking goals has been proven effective in enhancing students' critical thinking skills |
| Darmawan et al. [This research] | Development of Augmented Reality (AR) integrated e-content media to enhance students' critical thinking | The ARGEONET M6 is valid, practical, and effective, and enhances students' critical thinking |

Table 2 outlines previous studies concerning the integration of Metaverse and Augmented Reality (AR) technologies in mathematics education, as well as research on students' critical thinking in mathematical problem solving. Findings from Avci and AkgÜL (2024), Mutchima et al. (2025), and Rachmadtullah et al. (2023), demonstrate that the Metaverse fosters an immersive and interactive learning environment that enhances learning quality and supports the development of critical thinking. Studies by Chang et al. (2022), and Eşin and Özdemir (2022), reveal that AR-based learning can increase students' engagement and facilitate understanding of abstract mathematical concepts. Meanwhile, Khusna et al. (2024), Nosrat and Parsafar (2018), Wang and Abdullah (2024) emphasize that well-structured instructional designs incorporating problem-solving strategies, collaborative non-routine tasks, and clearly defined critical thinking objectives effectively strengthen students' higher-order thinking skills. Although several studies have examined the use of immersive technologies in education, there is still a lack of research focusing on the systematic development of Augmented Reality based e-content aimed at enhancing students' critical thinking skills. To address the research gap, this current study developed ARGEONET M6, an Augmented Reality integrated e-content designed to foster students' critical thinking in mathematics learning.

The ARGEONET design aligns with critical thinking indicators, namely focus, reason, inference, situation, clarity, and overview. First, to stimulate the focus indicator, ARGEONET provides instructional videos and AR visualizations that help students understand the relationship between the elements of three-dimensional shapes and their net. Second, the reason indicator is stimulated through justification activities to help students generate logical reasons and evidence for their answers. Third, the inference indicator is stimulated through investigation and problem-solving activities so that students can draw relevant conclusions. Fourth, the situation indicator is stimulated through tasks that require students to use appropriate mathematical information and concepts, such as in activities involving the identification of nets of flat-sided three-dimensional shapes. Fifth, the clarity indicator is stimulated through presentation and evaluation activities to provide justification for their thinking. Sixth, the overview indicator is stimulated through evaluation and reflection activities that guide students to review the solutions they have obtained. The aim of this study is to develop ARGEONET, an Augmented Reality integrated e-content designed to foster students' critical thinking in mathematics learning. Specifically, the study seeks to analyze the process and outcomes of developing ARGEONET to ensure that it meets the criteria of validity, practicality, and effectiveness to enhance students' critical thinking.

# 2. METHOD

## 2.1. Research Design

This research is a research and development (R&D) type of research employing the educational design framework proposed by Plomp (2013), which consists of three stages: (a) the preliminary research phase, (b) the prototyping phase, and (c) the assessment phase. In the preliminary research stage, a literature review, curriculum analysis, classroom observation, and test to determine students' critical thinking levels were conducted. Next, the prototype phase involves the development of instruments, testing the validity and reliability of the instruments, and developing the design and learning media of ARGEONET. The assessment stage is the final phase, in which an evaluation of the quality of the developed product ARGEONET, is carried out by conducting validity, practicality, and effectiveness tests based on the product quality criteria proposed by Nieveen and Folmer (2013). Validity was assessed through expert judgment, practicality was assessed through the implementation of the media and users' responses, and effectiveness was assessed through students' critical thinking achievement after using ARGEONET.

### 2.2. Participants and Research Setting

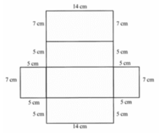
This research was conducted at a private junior high school located in Banyuwangi Regency, East Java, Indonesia. The school was chosen because its facilities and infrastructure support the implementation of technology-based learning. The selection was based on the criteria and specifications of devices for accessing ARGEONET. The first criterion is the availability of adequate internet access since ARGEONET is a web-based media that can only be accessed online. The second criterion is the availability of technological devices, whether school facilities or personal ownership by teachers and students. ARGEONET is compatible with Android, iOS, Windows, and macOS operating systems. Thus, ARGEONET can be accessed through various devices including smartphones, tablets, computers, or smart TVs equipped with browsers such as Google Chrome, Microsoft Edge, Mozilla Firefox, or the standard web browser available on the device.

Before the implementation is carried out, the researcher involved 15 students as respondents for the validity and reliability testing of the instrument. Next, this study involved 30 ninth-grade students of junior high school as a sample to conduct practicality and effectiveness tests of ARGEONET. The students involved in the validity and reliability testing of the instrument were at the same educational level and school but in different classes. The selection of the sample class employed purposive sampling based on teacher recommendations and considered the students' characteristics, specifically those who had acquired prerequisite knowledge on the elements of three-dimensional shapes. In addition, the students chosen as the sample had undergone learning treatment on three-dimensional nets using ARGEONET and completed pretest and posttest to analyze its effectiveness using a One Group Pretest-Posttest design. Therefore, the findings cannot yet be generalized to all junior high school students and are limited to students with similar characteristics and learning conditions. Furthermore, the assessment of ARGEONET's practicality was also conducted by a mathematics teacher who acted as the implementer. Meanwhile, the ARGEONET validation test by the researcher involved 2 experts in mathematics education who are lecturers with doctoral or professorial qualifications. To further explore the ARGEONET features that affect each component of critical thinking, the researchers interviewed 6 students who served as the study sample.

### 2.3. Research Instrument

The product validity instrument consists of validation sheets for ARGEONET learning media based on the Metaverse M6 learning model. These sheets were provided to two experts considered qualified to assess the aspects contained in the device or instrument being developed. Furthermore, the practicality instrument used in this study consisted of student and teacher user response questionnaires. These questionnaires were used to evaluate ease of use, the usefulness of features in facilitating student learning, the capability of the Metaverse M6-based ARGEONET media in enhancing students' understanding of three-dimensional nets, the attractiveness of design and interactivity, operational smoothness, relevance to the learning material, as well as recommendations. Meanwhile, the effectiveness of the ARGEONET media was measured using pretest and posttest instruments. The pretest and posttest instruments provided to the students are shown in Figure 1.

1. Salsa sedang menyusun jaring-jaring balok dari potongan kardus. Jaring-jaring yang dirancang Salsa memiliki ukuran sebagai berikut.

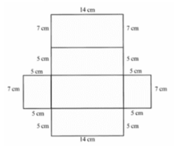



   Salsa menyusun balok tersebut dengan selotip sebagai perekatnya. Berdasarkan informasi-informasi yang diberikan, jawablah pertanyaan atau perintah berikut.
   a. Tuliskan informasi-informasi yang Anda pahami berdasarkan pernyataan dan gambar di atas!
   b. Berapakah panjang minimal selotip yang dibutuhkan oleh Salsa? Berikan alasan-alasan pada setiap langkah penyelesaian yang Anda lakukan!
   c. Kesimpulan apa yang Anda dapatkan berdasarkan jawaban soal b?
   d. Informasi apa saja yang Anda gunakan untuk menjawab soal b?
   e. Mengapa Anda membuat kesimpulan tersebut? (Beri alasan yang logis)
   f. Bagaimana cara Anda mengecek adanya kesalahan pada setiap langkah yang telah Anda lakukan? Jika tidak ada kesalahan, bagaimana cara Anda memastikan bahwa jawabanmu sudah benar?

2. Bu Rani berencana membuat souvenir miniatur piramida berbentuk limas segiempat yang kerangkanya terbuat dari seutas kawat. Alas piramida tersebut berbentuk persegi dengan ukuran rusuknya 6 $cm$, sedangkan seluruh rusuk tegaknya berukuran 8 $cm$. Berdasarkan informasi-informasi yang diberikan, jawablah pertanyaan atau perintah berikut.
   a. Tuliskan informasi-informasi yang Anda pahami berdasarkan pernyataan dan gambar di atas!
   b. Berapa panjang minimal kawat yang dibutuhkan Bu Rani untuk membuat kerangka miniatur piramida tersebut? Berikan alasan-alasan pada setiap langkah penyelesaian yang Anda lakukan!
   c. Kesimpulan apa yang Anda dapatkan berdasarkan jawaban soal b?
   d. Informasi apa saja yang Anda gunakan untuk menjawab soal b?
   e. Mengapa Anda membuat kesimpulan tersebut? (Beri alasan yang logis)
   f. Bagaimana cara Anda mengecek adanya kesalahan pada setiap langkah yang telah Anda lakukan? Jika tidak ada kesalahan, bagaimana cara Anda memastikan bahwa jawabanmu sudah benar?

**Translation**:

1. Salsa is arranging the net of a rectangular prism from pieces of cardboard. The net designed by Salsa has the following dimensions:

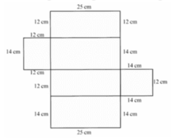



   Salsa arranged the blocks using adhesive tape as the bonding material. Based on the information provided, answer the following questions or instructions.
   a. Write down the information you understand based on the statement and the image above!
   b. What is the minimum length of tape required by Salsa? Provide reasons for each step of your solution!
   c. What conclusion can you draw based on your answer to question b?
   d. What pieces of information did you use to answer question b?
   e. Why did you reach that conclusion? (Provide a logical reason)
   f. How do you check for errors in each step you have taken? If there are no errors, how do you ensure that your answer is correct?

2. Ms. Rani plans to make pyramid miniature souvenirs in the form of a square pyramid whose frame is made of a piece of wire. The base of the pyramid is a square with side length of 6 cm, while all of its lateral edges are 8 cm long. Based on the information provided, answer the following questions or instructions.
   a. Write down the information you understand based on the statement and the image above!
   b. What is the minimum length of wire needed by Mrs. Rani to construct the pyramid miniature frame? Provide reasons for each step of your solution!
   c. What conclusion can you draw based on your answer to question b?
   d. What pieces of information did you use to answer question b?
   e. Why did you reach that conclusion? (Provide a logical reason)
   f. How do you check for errors in each step you have taken? If there are no errors, how do you ensure that your answer is correct?

(a)

1. Rina sedang membuat sebuah kotak penyimpanan berbentuk balok. Balok itu disusun dari beberapa potongan kardus dengan rancangan ukuran sebagai berikut.

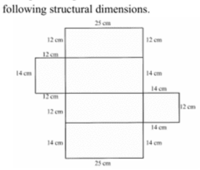



   Rina merekatkan potongan-potongan kardus tersebut dengan menggunakan lakban. Berdasarkan informasi-informasi yang diberikan, jawablah pertanyaan atau perintah berikut.
   a. Tuliskan informasi-informasi yang Anda pahami berdasarkan pernyataan dan gambar di atas!
   b. Berapakah panjang minimal lakban yang dibutuhkan oleh Rina? Berikan alasan-alasan pada setiap langkah penyelesaian yang Anda lakukan!
   c. Kesimpulan apa yang Anda dapatkan berdasarkan jawaban soal b?
   d. Informasi apa saja yang Anda gunakan untuk menjawab soal b?
   e. Mengapa Anda membuat kesimpulan tersebut? (Beri alasan yang logis)
   f. Bagaimana cara Anda mengecek adanya kesalahan pada setiap langkah yang telah Anda lakukan? Jika tidak ada kesalahan, bagaimana cara Anda memastikan bahwa jawabanmu sudah benar?

2. Tita berencana membuat miniatur tenda berbentuk limas segiempat yang kerangkanya terbuat dari lidi di mana tiap batangnya memiliki panjang 10 $cm$. Alas tenda tersebut berbentuk persegi dengan ukuran sisi 7 $cm$, sedangkan seluruh rusuk tegaknya berukuran 13 $cm$. Berdasarkan informasi-informasi yang diberikan, jawablah pertanyaan atau perintah berikut.
   a. Tuliskan informasi-informasi yang Anda pahami berdasarkan pernyataan dan gambar di atas!
   b. Berapakah banyak minimal batang lidi yang dibutuhkan oleh Tita? Berikan alasan-alasan pada setiap langkah penyelesaian yang Anda lakukan!
   c. Kesimpulan apa yang Anda dapatkan berdasarkan jawaban soal b?
   d. Informasi apa saja yang Anda gunakan untuk menjawab soal b?
   e. Mengapa Anda membuat kesimpulan tersebut? (Beri alasan yang logis)
   f. Bagaimana cara Anda mengecek adanya kesalahan pada setiap langkah yang telah Anda lakukan? Jika tidak ada kesalahan, bagaimana cara Anda memastikan bahwa jawabanmu sudah benar?

**Translation**:

1. Rina is making a block-shaped storage box. The block is assembled from several pieces of cardboard with the following structural dimensions.




   Rina seals the cardboard pieces together using adhesive tape. Based on the provided information, answer the following questions or instructions.
   a. Write down the information you understand based on the statement and the image above!
   b. What is the minimum length of tape required by Rina? Provide reasons for each step of your solution!
   c. What conclusion can you draw based on your answer to question b?
   d. What pieces of information did you use to answer question b?
   e. Why did you reach that conclusion? (Provide a logical reason)
   f. How do you check for errors in each step you have taken? If there are no errors, how do you ensure that your answer is correct?

2. plans to make a miniature tent in the shape of a rectangular pyramid, with a frame constructed from wooden sticks, where each stick has a length of 10 cm. The base of the tent is a square measuring 7 cm on each side, while all the lateral edges measure 13 cm. Based on the provided information, answer the following questions or instructions.
   a. Write down the information you understand based on the statement and the image above!
   b. What is the minimum number of wooden sticks required by Tita? Provide reasons for each step of your solution!
   c. What pieces of information did you use to answer question b?
   d. Why did you reach that conclusion? (Provide a logical reason)
   e. How do you check for errors in each step you have taken? If there are no errors, how do you ensure that your answer is correct?

(b)

**Figure 1**. (a) Pretest questions; (b) Posttest questions

The assessment of the pretest and posttest instruments in Figure 1 is based on the assessment rubric according to the critical thinking indicators in Table 1. Students' answers for each indicator are rated using a likert scale from 0 to 3 as shown in Table 3.

**Table 3**. Assessment rubric

| Critical Thinking Skills | Indicator | Score | Scoring Criteria |
|---|---|---|---|
| Focus | Writing down the necessary information to answer questions | 3 | Students write down all the information known in the question completely and accurately. |
| | | 2 | The students wrote down some of the information they knew on the question correctly, but there was still important information missing. |

| Critical Thinking Skills | Indicator | Score | Scoring Criteria |
|---|---|---|---|
| | | 1 | The student writes down some of the information they know from the question, but the information is inaccurate or doesn't match the problem. |
| | | 0 | The students didn't write down any information they knew in the question. |
| Reason | Providing reasoning based on relevant facts or evidence at every step of decision-making | 3 | The student uses all data and correct logical steps accompanied by clear reasons. |
| | | 2 | The student uses most of the data and correct logical steps, but the reasoning provided is incomplete or contains minor errors that do not affect the main result. |
| | | 1 | The student uses some relevant data, but the logical steps are not quite right, and the reasons given are unclear or still have some basic mistakes. |
| | | 0 | The students did not perform any calculations nor provided any reasoning. |
| Inference | Drawing correct conclusions based on appropriate reasoning | 3 | The student writes the conclusion correctly. |
| | | 2 | The student wrote the conclusion with the correct final answer but made an error in writing the units. |
| | | 1 | The student wrote a conclusion with an incorrect final answer. |
| | | 0 | The student did not write a conclusion. |
| Situation | Using all information according to the given problem | 3 | Students use information that is relevant to the problem appropriately. |
| | | 2 | The student uses some of the information that is relevant to the problem, but there is still information that is not utilized. |
| | | 1 | Students use information that is inaccurate or not relevant to the problem. |
| | | 0 | The student does not mention information that is relevant to the problem. |
| Clarity | Providing further explanation of the conclusions that have been made | 3 | Students write additional relevant explanations to reinforce the conclusion. |
| | | 2 | The student provides a general explanation and demonstrates a clear connection with the conclusions reached. |
| | | 1 | The student provides explanations that are general in nature but has not yet demonstrated a clear connection to the conclusions drawn. |
| | | 0 | The student did not provide additional explanations to reinforce the conclusion |
| Overview | Checking for errors that occur at each step that has been taken | 3 | The student writes the checking steps completely and logically. |
| | | 2 | The student conducted a check, but the steps written are incomplete. |
| | | 1 | The student stated that they had carried out the check but did not write down the steps or the steps were illogical. |
| | | 0 | The student did not perform the check. |

Effectiveness of pretest and posttest instruments is tested for reliability using the Intraclass Cronbach's Alpha (CA) and Correlation Coefficient (ICC). CA is used to test the consistency of test items, whereas ICC is used to assess scoring reliability. The criteria for the Intraclass Correlation Coefficient (ICC) test are presented in Table 4.

**Table 4**. The criteria for the Intraclass Correlation Coefficient (ICC)

| Score Rank | Criteria |
|---|---|
| $x < 0.50$ | Poor |
| $0.50 \leq x < 0.75$ | Moderate |
| $0.75 \leq x < 0.90$ | Good |
| $x \geq 0.90$ | Excellent |

(Source: Koo & Li, 2016)

In the reliability test, the results of students' pretest and posttest were assessed by two raters who hold bachelor's degrees in mathematics education and have more than two years of teaching experience at the junior high school level. Both raters used the same assessment rubric to ensure consistency and objectivity in the assessment process. The rubric was based on critical thinking indicators, mapping six types of indicators, there are focus, reason, inference, situation, clarity, and overview. Meanwhile, the criteria for the Cronbach's Alpha test are presented in Table 5.

**Table 5**. Cronbach's Alpha test criteria

| Cronbach's Alpha | Reliability |
|---|---|
| $CA < 0.70$ | Not Reliable |
| $CA \geq 0.70$ | Reliable |

(Source: Groenwold et al., 2018)

If the reliability test results prove the measurements are consistent, then the pretest and posttest instruments can be used. Furthermore, the students' scores after implementation were analyzed using a One Group Pretest–Posttest design. In addition, there is also a validated semi-structured interview guideline instrument to further explore the ARGEONET features that impact each component.

### 2.4. Analyzing of Data

#### 2.4.1. Validity of ARGEONET

Validators fill out assessment questionnaires on validation instruments containing item statements in accordance with the aspects observed for the purpose of tabulating data scores from the ARGEONET assessment. The ARGEONET validation scoring guidelines using a Likert scale ranging from 1 (very poor) to 5 (very good). The validator's assessment data is analyzed using a validity data analysis technique with the following formula.

$$Validity\ (V) = \frac{Total\ score}{Max\ score} \times 100\%$$

From the validity scores obtained, the researcher categorized the types of validity of the ARGEONET teaching media according to the validity criteria presented in Table 6.

**Table 6**. Validity criteria

| Score | Criteria |
|---|---|
| $85\% \leq V \leq 100\%$ | Very valid |
| $70\% \leq V < 85\%$ | Valid |
| $50\% \leq V < 70\%$ | Less valid |
| $1\% \leq V < 50\%$ | Not valid |

(Source: Kishore et al., 2021)

ARGEONET to enhance students' critical thinking is said to be valid if the validity percentage is $\geq 70\%$.

### 2.4.2. Practically of ARGEONET

In the practical test, data tabulation was collected from teachers and student's responses through filling out the ARGEONET teaching media practically test questionnaires. The scoring guidelines for ARGEONET practically test using a Likert scale ranging from 1 (very poor) to 5 (very good). The data obtained from the teacher responses were analyzed using practical data analysis techniques with the following formula.

$$Practicality\ percentage\ (P) = \frac{Total\ score}{Max\ Score} \times 100\%$$

From the average scores obtained, the researchers categorized the practicality of the ARGEONET teaching media according to the practicality test criteria presented in Table 7.

**Table 7**. Practically criteria

| Score | Criteria |
|---|---|
| $80\% \leq P \leq 100\%$ | Very practical |
| $60\% \leq P < 80\%$ | Practical |
| $40\% \leq P < 60\%$ | Quite Practical |
| $20\% \leq P < 40\%$ | Less practical |
| $0\% \leq P < 20\%$ | Impractical |

(Source: Rassyi et al., 2023)

ARGEONET to enhance students' critical thinking is considered practical if the *P* percentage is $\geq 61\%$.

### 2.4.3. Effectiveness of ARGEONET

The effectiveness of ARGEONET can be obtained from student learning outcomes after using ARGEONET. Students' learning outcomes are measured through pretests and posttest scores. The data were analyzed quantitatively using a One Group Pretest–Posttest design with SPSS, following these steps.

a. The researchers collected pretest and posttest scores
b. Data collected according to the aspects of critical thinking are gradually entered into the software for further analysis.
c. The researchers tested the normality of the data on the difference in scores between the pretest and posttest for each aspect using the Shapiro–Wilk test. If the significance value is greater than 0.05,

the data is considered normal. Conversely, if the significance value is less than or equal to 0.05, the data is considered non-normal.

d. The researchers continued their analysis based on the results of the normality test. Normally distributed data were analyzed using a Paired Sample t-test, while non-normally distributed data were analyzed using the Wilcoxon Signed Rank Test.
e. In each test, the researcher applied a significance level of $\alpha = 0{,}05$
f. This study examines whether there are differences in the aspects of critical thinking, including focus, reason, inference, situation, clarity, and overview, observed in students. The following hypotheses were tested.
   $H_0$: there is no difference in student's critical thinking across the aspects examined ($p > 0.05$).
   $H_1$: there is difference in student's critical thinking across the aspects examined ($p \leq 0.05$).

After analyzing the differences in students' critical thinking in each aspect, the researcher conducted calculations using Wilcoxon Effect Size Test. These calculations were carried out to determine the effect size of Metaverse-based ARGEONET M6 on students' critical thinking. The formula for the Wilcoxon effect size test is presented as follows.

$$r = \frac{|Z|}{\sqrt{n}}$$

In that formula, r is the effect size value, Z is the Wilcoxon test statistic value, while n is the number of samples or data pairs analyzed. From the value obtained, the researchers categorized the effect size of the ARGEONET teaching media according to the criteria presented in Table 8.

**Table 8**. Wilcoxon effect size test criteria

| Range of *r* Values | Effect Size |
|---|---|
| $r < 0.30$ | Small |
| $0.30 \leq r < 0.50$ | Medium |
| $0.50 \leq r$ | Large |

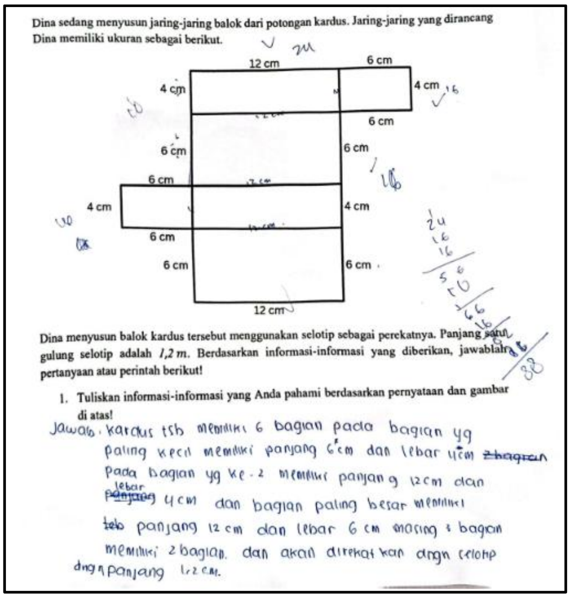
(Source: Cohen, 1992)

## 3. RESULTS AND DISCUSSION

### 3.1. Results

#### 3.1.1. Preliminary Phase

Ennis (1996) states that critical thinking consists of six components: (1) focus, (2) reason, (3) inference, (4) situation, (5) clarity, and (6) overview. In the preliminary study, certain critical thinking indicators were not achieved by students in solving mathematical problems. The first was the focus indicator, which students had not achieved, as shown in Figure 2.


Dina sedang menyusun jaring-jaring balok dari potongan kardus. Jaring-jaring yang dirancang Dina memiliki ukuran sebagai berikut.

Dina menyusun balok kardus tersebut menggunakan selotip sebagai perekatnya. Panjang satu gulung selotip adalah 1,2 m. Berdasarkan informasi-informasi yang diberikan, jawablah pertanyaan atau perintah berikut!

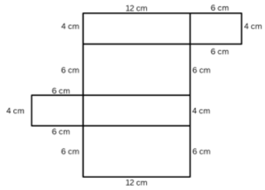
1. Tuliskan informasi-informasi yang Anda pahami berdasarkan pernyataan dan gambar di atas!

Jawab: kardus tsb memiliki 6 bagian pada bagian yg paling kecil memiliki panjang 6 cm dan lebar 4 cm ~~2 bagian~~ pada bagian yg ke-2 memiliki panjang 12 cm dan lebar 4 cm dan bagian paling besar memiliki panjang 12 cm dan lebar 6 cm masing² bagian memiliki 2 bagian. dan akan direkatkan dgn selotip dgn panjang 1,2 cm.



**Translation:**

Dina is arranging nets of a rectangular prism from pieces of cardboard. The nets designed by Dina have the following dimensions.



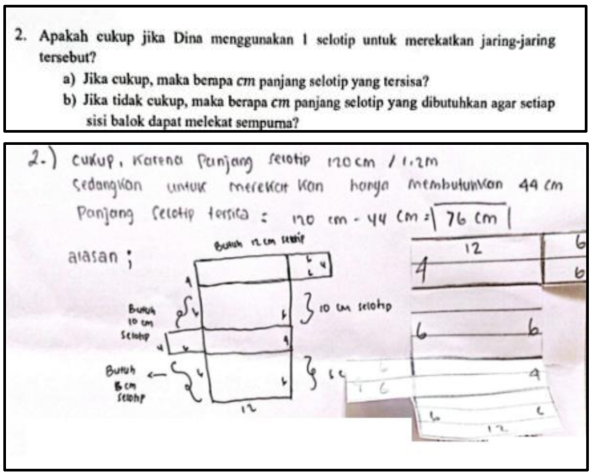

Dina arranged the cardboard blocks using tape as an adhesive. The length of one roll of tape is 1.2 m. Based on the information provided, answer the following questions or instructions!

1. Write down the information you understand from the statements and the picture above!

Answer: The cardboard has 6 sections; the smallest section has a length of 6 cm and a width of 4 cm. The second section has a length of 12 cm and a width of 4 cm, and the largest section has a length of 12 cm and a width of 6 cm. Each section has 2 parts and will be taped with a length of 1.2 cm.

**Figure 2**. Student A's answer indicating the unachievement of the focus indicator

Figure 2 shows that the student has not achieved the focus indicator. This is because the student was unable to write down the necessary information to answer the question, which included the length, width, and height of the cuboid. Instead, the student presented the dimensions of a rectangle, namely length and width, which could potentially lead to misconceptions when used to solve the problem. An example of a misconception experienced by the student due to using irrelevant data to solve the problem is presented in Figure 3.


2. Apakah cukup jika Dina menggunakan 1 selotip untuk merekatkan jaring-jaring tersebut?
   a) Jika cukup, maka berapa cm panjang selotip yang tersisa?
   b) Jika tidak cukup, maka berapa cm panjang selotip yang dibutuhkan agar setiap sisi balok dapat melekat sempurna?

2.) cukup, karena panjang selotip 120 cm / 1,2 m sedangkan untuk merekatkan hanya membutuhkan 44 cm
Panjang selotip tersisa : 120 cm - 44 cm = 76 cm
alasan ;

**Translation:**

1. Is it enough if Dina uses 1 piece of tape to attach the netting?
   a) If it is enough, how many centimeters of tape are left?
   b) If it is not enough, how many centimeters of tape are needed so that each side of the block can be attached perfectly?

Answer:

It is enough because the tape is 120 cm/1.2 m long, whereas attaching only requires 44 cm. Remaining tape length = 120 cm - 44 cm = 76 cm. Reason:

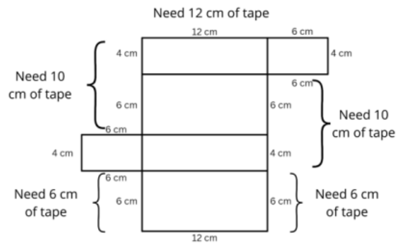
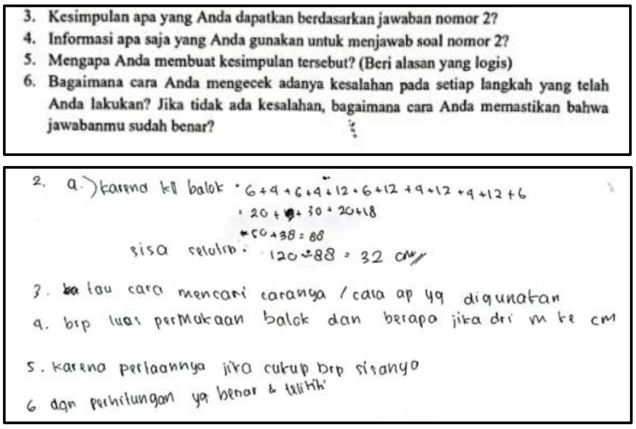



**Figure 3**. Student's misconception

In Figure 3, the student experienced a misconception by concluding that the tape needed to attach the sides of the cuboid completely was 44 cm. The data used by the student were related to the length and width of a rectangle. However, the correct problem-solving process requires the use of length, width, and height of the cuboid, which results in 88 cm of tape being needed to attach the sides properly. Therefore, the focus indicator is important for students to achieve in order to avoid further misconceptions.

The second critical thinking indicator that students failed to achieve is the situation indicator. Evidence of the unachievement of the situation indicator by students is presented in Figure 4.


3. Kesimpulan apa yang Anda dapatkan berdasarkan jawaban nomor 2?
4. Informasi apa saja yang Anda gunakan untuk menjawab soal nomor 2?
5. Mengapa Anda membuat kesimpulan tersebut? (Beri alasan yang logis)
6. Bagaimana cara Anda mengecek adanya kesalahan pada setiap langkah yang telah Anda lakukan? Jika tidak ada kesalahan, bagaimana cara Anda memastikan bahwa jawabanmu sudah benar?

2. a.) karena kll balok = 6+4+6+4+12+6+12+4+12+4+12+6
= 20 + 10 + 30 + 20+18
= 50 + 38 = 88
sisa seluruh = 120 − 88 = 32 cm
3. tau cara mencari caranya / cara ap yg digunakan
4. brp luas permukaan balok dan berapa jika dri m ke cm
5. karena pertaannya jika cukup brp sisanya
6 dan perhitungan yg benar & teliti

**Translation:**

3. What conclusion do you draw based on the answer to question number 2?
4. What information did you use to answer question number 2?
5. Why did you make that conclusion? (Provide a logical reason)
6. How do you check for errors in each step you have taken? If there are no errors, how do you ensure that your answer is correct?

**Answer:**

2. a) because of the perimeter of the cuboid $= 6 + 4 + 6 + 4 + 12 + 6 + 12 + 4 + 12 + 4 + 12 + 6$
$= 20 + 30 + 20 + 18$
$= 50 + 38 = 88$
remaining tape $= 120 - 88 = 32\ cm$
3. Know the method to find the way/which method is used.
4. Determine the surface area of the rectangular prism and convert it from meters to centimeters if needed.
5. Because the question asks how much is sufficient, determine the remainder.
6. Perform the calculations accurately and meticulously.

**Figure 4**. Student B’s answer indicating the unachievement of the situation indicator

Figure 4. shows that the student was unable to achieve the situation indicator. This is because the student could not use all the information according to the given problem. In Figure 4, the student wrote that the information used to answer question number 2 was the surface area of the cuboid and the unit conversion from meters to centimeters. In fact, the information needed to answer question number 2 included the lengths of the cuboid’s edges (length, width, and height), the length of the tape, and the calculation of the tape required to attach the net.

The third critical thinking indicator that students failed to achieve is the clarity indicator. Evidence of the unachievement of the clarity indicator is presented in Figure 5.


5.) karena memang kesimpulannya adalah itu, bagian akhir dari persoalan.

**Translation**:
5. because indeed the conclusion is that, the final part of the matter

**Figure 5**. Student C’s answer indicating the unachievement of the clarity indicator

In Figure 5, the student wrote that the conclusion was the final part of the problem. This shows that the student was unable to state the reasoning behind the conclusion provided, indicating that the clarity indicator was not achieved. The unachievement of certain critical thinking indicators by students demonstrates the need for stimuli to enhance their critical thinking skills. One possible stimulus is the implementation of the MathVerse M6 learning model assisted by ARGEONET to improve critical thinking ability.

### 3.1.2. Prototyping Phase

During the prototype phase, prototyping was conducted on two types of research instruments, there are the pretest and posttest instruments and the ARGEONET platform. The objective was to develop pretest and posttest instruments to measure improvements in students’ critical thinking skills. Furthermore, the prototype phase was also conducted to produce a functional ARGEONET prototype that could be evaluated in terms of pedagogical, content, technology, visual design and multimedia, engagement and motivation, and accessibility and affordability aspects.

### 3.1.2.1. Prototyping pretest and posttest Instruments

During the prototype phase of the pretest and posttest, validity and reliability tests were conducted. The validity test was conducted by experts, and an empirical validity test was performed using Pearson Product-moment Correlation. Meanwhile, the reliability test utilized the ICC and Cronbach's Alpha.

***Instrument Validation***

Before being used in the study, pretest and posttest measures of mathematical critical thinking skills were pilot-tested on students with characteristics similar to those of the study participants. Validation was conducted in two ways, there are expert validation by a doctor in mathematics education and empirical validation using Pearson's correlation coefficient, which involved 15 respondents. First, the results of the expert validity test are presented in Table 9.

**Table 9**. Results of the expert validity test on the test instrument

| Test Instrument | Aspect | Average Score | |
|---|---|---|---|
| | | Validator 1 | Validator 2 |
| Pretest | Material | 3.67 | 2.67 |
| | Language | 3 | 4 |
| | Question Construction | 3.5 | 4 |
| | Total Average | 3.39 (84.75%) | 3.56 (89%) |
| Posttest | Material | 3.67 | 2.67 |
| | Language | 3 | 4 |
| | Question Construction | 3.5 | 4 |
| | Total Average | 3.39 (84.75%) | 3.56 (89%) |

Second, item validity was analyzed using Pearson's product–moment correlation. An item was deemed valid if the calculated correlation coefficient ($r$-test) was greater than the critical value of Pearson's correlation coefficient ($r$-table) at the selected significance level ($\alpha$=0.05), adjusted for sample size. The results of the empirical validation of the pretest items are presented in Table 10.

**Table 10**. Results of the Pearson product-moment correlation validity test

| No | Test Instrument | Critical Thinking Indicator | $r$-test | Validity |
|---|---|---|---|---|
| 1 | Pretest | Focus | Question 1: 0.962 > 0.514 | Question 1: Valid |
| | | | Question 2: 0.760 > 0.514 | Question 2: Valid |
| | | Reason | Question 1: 0.910 > 0.514 | Question 1: Valid |
| | | | Question 2: 0.936 > 0.514 | Question 2: Valid |
| | | Inference | Question 1: 0.922 > 0.514 | Question 1: Valid |
| | | | Question 2: 0.946 > 0.514 | Question 2: Valid |

| No | Test Instrument | Critical Thinking Indicator | *r*-test | Validity |
|---|---|---|---|---|
| | | Situation | Question 1: 0.961 > 0.514<br>Question 2: 0.952 > 0.514 | Question 1: Valid<br>Question 2: Valid |
| | | Clarity | Question 1: 1.000 > 0.514<br>Question 2: 1.000 > 0.514 | Question 1: Valid<br>Question 2: Valid |
| | | Overview | Question 1: 0.967 > 0.514<br>Question 2: 0.965 > 0.514 | Question 1: Valid<br>Question 2: Valid |
| 2 | Posttest | Focus | Question 1: 0.882 > 0.514<br>Question 2: 0.932 > 0.514 | Question 1: Valid<br>Question 2: Valid |
| | | Reason | Question 1: 0.968 > 0.514<br>Question 2: 0.955 > 0.514 | Question 1: Valid<br>Question 2: Valid |
| | | Inference | Question 1: 0.966 > 0.514<br>Question 2: 0.963 > 0.514 | Question 1: Valid<br>Question 2: Valid |
| | | Situation | Question 1: 0.919 > 0.514<br>Question 2: 0.838 > 0.514 | Question 1: Valid<br>Question 2: Valid |
| | | Clarity | Question 1: 0.952 > 0.514<br>Question 2: 0.910 > 0.514 | Question 1: Valid<br>Question 2: Valid |
| | | Overview | Question 1: 0.976 > 0.514<br>Question 2: 0.933 > 0.514 | Question 1: Valid<br>Question 2: Valid |

Table 10 shows that each critical thinking indicator in the pretest and posttest items was found to be valid because all r-test values for each item were greater than the r-table value (0.514). Thus, all test items met the criteria for empirical validity and could be used to measure students' critical thinking skills during the implementation phase of the study.

***Instrument Reliability Test***

Reliability testing was conducted using the Intraclass Correlation Coefficient (ICC) and Cronbach's Alpha. The ICC was used to calculate inter-rater reliability, which measures the level of consistency in scoring between two raters regarding students' responses to open-ended questions. The ICC was chosen because the research instrument consists of essay questions that require a subjective scoring process, thus necessitating a test of inter-rater agreement. Meanwhile, Cronbach's Alpha was used to test the instrument's internal consistency, that is the extent to which the two essay items consistently measure the same construct. Cronbach's Alpha was chosen because the instrument consists of two items designed to measure the same aspect, thus requiring information regarding the relationship and consistency between items. Therefore, the combination of ICC and Cronbach's Alpha can measure the instrument's reliability in terms of both inter-rater assessment consistency and item-internal consistency.

The ICC test results for the pretest and posttest instruments from the raters for each critical thinking indicator are presented in Table 11.

**Table 11**. Results of the instrument reliability test

| No | Instrument Type | Critical Thinking Indicator | ICC score | Reliability |
|---|---|---|---|---|
| 1 | Pretest | Focus | 0.909 | Excellent |
| | | Reason | 1.000 | Excellent |
| | | Inference | 1.000 | Excellent |
| | | Situation | 0.812 | Good |
| | | Clarity | 1.000 | Excellent |
| | | Overview | 1.000 | Excellent |
| 2 | Posttest | Focus | 0.937 | Excellent |
| | | Reason | 0.988 | Excellent |
| | | Inference | 0.989 | Excellent |
| | | Situation | 0.855 | Good |
| | | Clarity | 0.965 | Excellent |
| | | Overview | 0.945 | Excellent |

Table 11 shows that the Intraclass Correlation Coefficient (ICC) values for the average measures on the pretest and posttest for the indicators of focus, reason, inference, clarity, and overview meet the criteria for the "excellent" category. Meanwhile, in the pretest, the situation indicator met the "good" criteria. This indicates that the inter-rater scoring results exhibit a high level of consistency and agreement, meaning the assessment instrument is capable of producing reliable measurements and can be used in research. Furthermore, the results of the Cronbach's Alpha test are presented in Table 12.

**Table 12**. Result of the Cronbach's Alpha

| No | Instrument Type | Critical Thinking Indicator | Cronbach's Alpha | Reliability |
|---|---|---|---|---|
| 1 | Pretest | Focus | 0.909 > 0.70 | Reliable |
| | | Reason | 1.000 > 0.70 | Reliable |
| | | Inference | 1.000 > 0.70 | Reliable |
| | | Situation | 0.812 > 0.70 | Reliable |
| | | Clarity | 1.000 > 0.70 | Reliable |
| | | Overview | 1.000 > 0.70 | Reliable |
| 2 | Posttest | Focus | 0.937 > 0.70 | Reliable |
| | | Reason | 0.988 > 0.70 | Reliable |
| | | Inference | 0.989 > 0.70 | Reliable |
| | | Situation | 0.855 > 0.70 | Reliable |
| | | Clarity | 0.965 > 0.70 | Reliable |
| | | Overview | 0.945 > 0.70 | Reliable |

Table 12 shows that each critical thinking indicator in the pretest and posttest instruments met the reliability criteria, as evidenced by Cronbach's Alpha values greater than 0.70. This indicates that each indicator demonstrated a high level of consistency in measuring students' critical thinking skills. Therefore, the pretest and posttest instruments were deemed reliable and suitable for use as data collection tools in this study.

### 3.1.2.2. Prototyping ARGEONET

ARGEONET is a technology-based learning media for three-dimensional geometric nets that integrates with Metaverse and M6 learning model. Therefore, ARGEONET is designed in accordance with the M6 syntax. The ARGEONET home display and the learning content menu is presented in Figure 6.


(a) (b)

**Figure 6**. (a) ARGEONET home display; (b) Learning content menu display

In accordance with the sequence and the number of M6 syntax, there are 6 learning content menus in the ARGEONET media. The first menu, *Activity 1: Let's Focus on Initial Ability*, presents an instructional video on the nets of three-dimensional shapes accompanied by augmented reality visualization. The second menu, *Activity 2: Let's Justify*, contains learning videos illustrating examples and non-examples of flat-faced 3D shape nets, also supported by augmented reality. The display of the learning video and augmented reality available in this media is shown in Figure 7.


(a) (b)

**Figure 7**. (a) Learning video; (b) 3D model based on augmented reality

In addition to the menu presenting the two learning e-contents shown in Figure 7, ARGEONET also includes a learning menu containing problems to enhance students' critical thinking in the third menu. The third menu, *Activity 3: Let's Investigate*, investigate engages students in critical thinking processes by presenting problems that require focusing questions and identifying the necessary information, analyzing the reasons for decision-making, formulating conclusions, and providing further explanations of the conclusions reached. The provision of problems that require critical thinking is also given in the *Competency Test menu*.

This menu allows students to assess their understanding through an interactive quiz that provides real-time scoring feedback. The display of the investigation and competency test menu is shown in Figure 8.


(a) (b)

**Figure 8**. (a) Problems in the investigation menu; (b) Competency test

After students complete Activities 1 to 3 as shown in 8.a, they can evaluate their work by accessing the fourth and fifth menus. The fourth menu, *Activity 4: Let's Present Ideas*, encourages students to summarize their responses from the previous activities. The fifth menu, *Activity 5: Let's Evaluate Answers*, prompts students to critically assess and justify the outcomes of their group discussions. The summary display in activity 4 and the instructions from activity 5 are shown in Figure 9.


(a) (b)

**Figure 9**. (a) Summary of activities presented in Activity 4; (b) Evaluation guidelines in Activity 5

After presenting and evaluating the answers, ARGEONET directs to the sixth menu, *Activity 6: Let's Conclude Answers*, requires students to formulate and write their final conclusions. In addition, ARGEONET also includes access buttons for the user guide, learning objectives, the button to print and share answers, and information about the ARGEONET application itself. Overall, the seven primary features integrated into ARGEONET represent an implementation of the M6 learning model integrated with the Metaverse (MathVerse M6), designed to promote the development of students' critical thinking skills.

### 3.1.3. Assessment Phase

The third phase is the assessment stage, where research evaluated whether the validator prototype met the criteria being valid, practical, and effective, so determined its feasibility for use.

#### 3.1.3.1. Validity Test

The ARGEONET media was validated by two expert validators, namely doctoral-level mathematics education lectures. The validation results from Validator 1 and Validator 2 are presented in Table 13.

**Table 13**. ARGEONET validation results by Validator 1 and Validator 2

| Aspect | Average Score | |
|---|---|---|
| | Validator 1 | Validator 2 |
| Pedagogical aspect (learning effectiveness) | 4.5 | 4.5 |
| Content aspect (material accuracy) | 4 | 4.75 |
| Technology aspect (performance and functionality) | 4.5 | 4.75 |
| Visual design and multimedia aspect | 4.75 | 4.5 |
| Engagement and motivation aspect | 4.67 | 4.67 |
| Accessibility and affordability aspect | 4.25 | 4.75 |
| Average | 4.45 (89%) | 4.65 (93%) |

Table 13 shows that the validation result of ARGEONET by Validator 1 obtained an overall average score of 4.45 out of 5.00, or equivalent to 89%. Thus, ARGEONET falls into the 'very valid' category. Furthermore, the validation results of ARGEONET by Validator 2 obtained an overall average score of 4.65 out 5.00, or equivalent to 93%, which is categorized as 'very valid'. Based on the validation result from both validators, the ARGEONET media is declared feasible for use and can be tested in the learning process.

#### 3.1.3.2. Practically Test

The practicality assessment of the media was carried out by grade IX students and teachers. The practicality instrument used was a user response questionnaire for both students and teachers with a 1–5 Likert scale. The response questionnaire consisted of seven statements: (1) easy to use, (2) features that facilitate student learning, (3) ARGEONET can improve students' understanding of net of 3D shapes, (4) attractiveness of design and interactivity, (5) smooth operation of the website including the presence of bugs and errors in use, (6) relevance or appropriateness of the material, and (7) users' recommendation of the website to others. ARGEONET was considered practical if the percentage reached $\geq$ 61%. Meanwhile, the results of the recapitulation of practicality based on the user response scores are presented in Table 14.

**Table 14**. Recapitulation of student and teacher response questionnaire scores

| Aspect Statement | Average Score | |
|---|---|---|
| | **Student Responses** | **Teacher Responses** |
| Easy to use | 4.60 | 5.00 |
| Features facilitate student learning | 4.32 | 5.00 |
| ARGEONET enhances students' understanding of nets of 3D shapes | 4.48 | 5.00 |
| Attractiveness of design and interactivity | 4.24 | 3.00 |
| Smooth operation of the website | 4.20 | 5.00 |
| Relevance or appropriateness of the material | 4.28 | 5.00 |
| Users recommend the website to others | 4.00 | 4.00 |
| Mean | 4.30 | 4.57 |
| Percentage | 86.06% | 91.43% |
| Category | Very practical | Very practical |

Table 14 illustrates that the percentage of the practicality score of the ARGEONET media obtained through the student and teacher response questionnaires reached 86.06% and 91.43%. This percentage indicates that ARGEONET is classified in the very practical category. This demonstrates that the ARGEONET media developed has fulfilled the feasibility criteria in terms of practicality. Aspects such as ease of use, smooth operation, material relevance, and the media's ability to enhance understanding of 3D shape nets received high ratings, indicating that ARGEONET is easy to use and effectively supports learning. Although the design and interactivity aspects received slightly lower scores from teachers, overall, the media falls into the "very practical" category and is suitable for use in mathematics learning.

#### 3.1.3.3. Effectiveness Test of ARGEONET on Students' Critical Thinking

***Focus***

This section presents the results of a test of ARGEONET's effectiveness in improving the focus indicator. The first step in this analysis is to conduct a descriptive statistical analysis to determine the distribution of scores and assess student performance before and after the intervention. Table 15 presents descriptive statistics for students' pretest and posttest scores in the focus indicator.

**Table 15**. Descriptive statistics of pretest and posttest scores in the focus indicator

| Types of Tests | | Statistic | Std. Error |
|---|---|---|---|
| Pretest_Focus | Mean | 4.8000 | 0.27709 |
| | Std. Deviation | 1.51771 | |
| | Minimum | 2.00 | |
| | Maximum | 6.00 | |
| Posttest_Focus | Mean | 5.566667 | 0.170754 |
| | Std. Deviation | 0.935261 | |
| | Minimum | 2.00 | |
| | Maximum | 6.00 | |

Table 15 presents descriptive statistics for students' pretest and posttest scores, showing a significant improvement in overall performance following the implementation of the ARGEONET medium. The mean score increased from 4.8 on the pretest to 5.67 on the posttest. These findings indicate an improvement in the indicator of students' focus on critical thinking. The standard deviation decreased slightly from 1.51771 to 0.935261, indicating reduced score variability and a more consistent level of achievement among participants following the use of ARGEONET. A one-group pretest-posttest design was used. First, a normality test was conducted on students' pretest and posttest responses for the focus indicator presented in Table 16.

**Table 16**. Test of normality

| Types of Test | Shapiro-Wilk | | |
|---|---|---|---|
| | Statistic | df | Sig. |
| Focus_Pretest | 0.743 | 30 | 0.000 |
| Focus_Posttest | 0.539 | 30 | 0.000 |

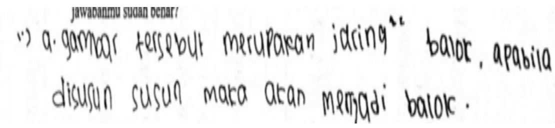
Table 16 shows that the results of the normality test using the Shapiro–Wilk test indicate a significance value (Sig.) of 0.000 for the focus indicator in both the pretest and posttest. Both significance values are less than the threshold of α = 0.05 (Sig. < 0.05), indicating that the data in both groups are not normally distributed. Because the data are not normally distributed, a nonparametric test specifically, the Wilcoxon test. The results of the Wilcoxon test are presented in Table 17.

**Table 17**. Results of the Wilcoxon test

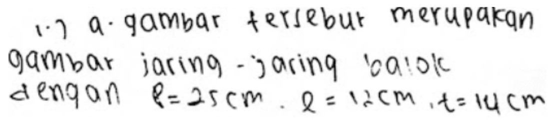
| | Focus_Posttest - Focus_Pretest |
|---|---|
| Z | -2.871 |
| Asymp. Sig. (2-tailed) | 0.004 |

Table 17 shows that the significance value (Sig.) of the Wilcoxon test is 0.004 for the focus indicator. This significance indicates that there are differences in students' critical thinking skills regarding the focus indicator. Furthermore, an effect size test was conducted to determine the extent to which the use of ARGEONET improved the achievement of the focus indicators. The effect size test revealed that $r = \frac{|Z|}{\sqrt{n}} = \frac{|-2.871|}{\sqrt{30}} = 0.524 > 0.50$. Therefore, the impact of using ARGEONET is large to improve the focus indicator of critical thinking.

Next, excerpts of pretest and posttest answers are presented, as well as excerpts from in-depth interviews with students who experienced an increase in scores on the focus indicator in critical thinking skills. One of the subjects who experienced an increase in score on the focus indicator, referred to by the researcher as S1. Figure 10 shows S1's answers in the pretest and posttest.




**Translation**:
That image is a net of a rectangular prism, when assembled, it will form a rectangular prism.
(a)




**Translation**:
That image is a net of a rectangular prism with a length of 25 cm, a width of 12 cm, and a height of 14 cm.
(b)

**Figure 10**. (a) S1' pretest answer related to the focus indicator; (b) S1' posttest answers related to the focus indicator

Figure 10(a) shows S1's answer before the implementation of ARGEONET in learning. S1 only wrote that the nets provided could form a rectangular prism. Meanwhile, Figure 10(b) shows S1's answer after the implementation of ARGEONET. S1 wrote complete information, starting from the nets in the question, which are nets of a rectangular prism, to the length, width, and height of the resulting rectangular prism. This explanation is supported by S1's statement in Interview 1.

*Interview 1*

*R : Do you feel like your answers on the pretest and posttest are different?*

*S1 : Hmm, if I look again, it seems like the pretest answers aren't as complete as the posttest ones.*

*R : Why did your way of answering differ between the two tests?*

*S1 : I only knew that it was the net of a rectangular prism. I hadn't thought about writing down the dimensions because I was just focusing on looking at the shape. But after that, I could see the parts of the rectangular prism more clearly, especially when using the media, so I knew the length, width, and height of the 3D shape.*

Based on interview 1, S1 said that after using ARGEONET, S1 felt it was easier to observe each part of the block net and understand the information contained in the image. Thus, this proves an improvement in S1's ability to identify and write down the information known from the problem more completely and accurately.

***Reason***

This section presents the results of a test of ARGEONET's effectiveness in improving the reason indicator. The first step in this analysis is to conduct a descriptive statistical analysis to determine the distribution of scores and assess student performance before and after the intervention. Table 18 presents descriptive statistics for students' pretest and posttest scores in the reason indicator.

**Table 18**. Descriptive statistics of pretest and posttest scores in the reason indicator

| Types of Tests | | Statistic | Std. Error |
|---|---|---|---|
| Pretest_Reason | Mean | 2.433333 | 0.132902 |
| | Std. Deviation | 0.727932 | |
| | Minimum | 2.00 | |
| | Maximum | 5.00 | |

| Types of Tests | | Statistic | Std. Error |
|---|---|---|---|
| Posttest_Reason | Mean | 3.633333 | 0.277440 |
| | Std. Deviation | 1.519604 | |
| | Minimum | 2.00 | |
| | Maximum | 6.00 | |

Table 18 presents descriptive statistics for students' pretest and posttest scores, showing a significant improvement in overall performance following the implementation of the ARGEONET medium. The mean score increased from 2.433 on the pretest to 3.633 on the posttest. These findings indicate an improvement in the indicator of students' reason on critical thinking. The standard deviation increased from 0.727932 to 1.519604, indicating a wider distribution of performance scores among participants following the use of ARGEONET. A one-group pretest-posttest design was used. First, a normality test was conducted on students' pretest and posttest responses for the reason indicator presented in Table 19.

**Table 19**. Test of normality

| Types of Test | Shapiro-Wilk | | |
|---|---|---|---|
| | **Statistic** | **df** | **Sig.** |
| Reason_Pretest | 0.637 | 30 | 0.000 |
| Reason_Posttest | 0.828 | 30 | 0.000 |

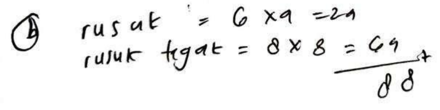
Table 19 shows that the results of the normality test using the Shapiro–Wilk test indicate a significance value (Sig.) of 0.000 for the reason indicator in both the pretest and posttest. Both significance values are less than the threshold of $\alpha = 0.05$ (Sig. < 0.05), indicating that the data in both groups are not normally distributed. Because the data are not normally distributed, a nonparametric test specifically, the Wilcoxon test. The results of the Wilcoxon test are presented in Table 20.

**Table 20**. Results of the Wilcoxon test

| | Reason_Posttest - Reason_Pretest |
|---|---|
| Z | -3.442 |
| Asymp. Sig. (2-tailed) | 0.001 |

Table 20 shows that the significance value (Sig.) of the Wilcoxon test is 0.001 for the reason indicator. This significance indicates that there are differences in students' critical thinking skills regarding the reason indicator. Furthermore, an effect size test was conducted to determine the extent to which the use of ARGEONET improved the achievement of the reason indicators. The effect size test revealed that $r = \frac{|Z|}{\sqrt{n}} = \frac{|-3.442|}{\sqrt{30}} = 0.628 > 0.50$. Therefore, the impact of using ARGEONET is large to improve the reason indicator of critical thinking.

Next, excerpts of pretest and posttest answers are presented, as well as excerpts from in-depth interviews with students who experienced an increase in scores on the reason indicator in critical thinking skills. One of the subjects who experienced an increase in score on the reason indicator, referred to by the researcher as S2. Figure 11 shows S2's answers in the pretest and posttest.




**Translation:**

Edges = 6 × 4 = 24
Lateral edges= 8 × 8 = 64 +
88

(a)


**Translation:**

7 × 4 = 28
13 × 4 = 52 +
80
↓
The total length can be obtained from the sum of all its edges

because 1 stick is 10cm, many sticks
$\frac{80}{10}$ =8 stick

(b)

**Figure 11**. (a) S2' pretest answer related to the reason indicator; (b) S2' posttest answers related to the reason indicator

Figure 11(a) shows S2's answer before the implementation of ARGEONET in learning. S2 made a mistake in determining the strategy. The calculation that should have resulted in a final answer of 56 cm, because of the wrong strategy, became 88 cm. In addition, S2 also did not provide reasons why the strategies and answers were written that way. On the other hand, Figure 11(b) shows S2's answers after using ARGEONET. S2 used all existing data or information and the correct logical steps accompanied by clear reasons. A more detailed explanation regarding the student's answers is presented in the excerpt of Interview 2.

*Interview 2*

*R : Why did you use that kind of solution step on the pretest and posttest questions?*

*S2 : Because what is being sought is the total edge length of a square pyramid, the length of the base edges plus the length of the lateral edges.*

*R : If so, why are the steps you used different between the pretest and posttest answers?*

*S2 : In the pretest, I still did not understand the shape of the figure, so I directly calculated based on the parts visible in the image. After using ARGEONET, I better understood the arrangement of the edges on the pyramid and realized that all edges must be counted with the correct arrangement, namely 4 base edges and 4 vertical edges.*

Based on interview 2, S2 mentions understanding the arrangement of the edges on a pyramid and knowing that all edges must be counted with their exact arrangement, namely 4 base edges and 4 vertical edges through the use of ARGEONET. Thus, this indicates an improvement in the critical thinking skills of the S2 on the reason indicator, where S2 can use all data and logical steps correctly accompanied by clear reasoning.

### *Inference*

This section presents the results of a test of ARGEONET's effectiveness in improving the inference indicator. The first step in this analysis is to conduct a descriptive statistical analysis to determine the distribution of scores and assess student performance before and after the intervention. Table 21 presents descriptive statistics for students' pretest and posttest scores in the inference indicator.

**Table 21**. Descriptive statistics of pretest and posttest scores in the inference indicator

| Types of Tests | | Statistic | Std. Error |
|---|---|---|---|
| Pretest_Inference | Mean | 2.700000 | 0.192682 |
| | Std. Deviation | 1.055364 | |
| | Minimum | 2.00 | |
| | Maximum | 5.00 | |
| Posttest_Inference | Mean | 3.633333 | 0.269027 |
| | Std. Deviation | 1.473521 | |
| | Minimum | 2.00 | |
| | Maximum | 6.00 | |

Table 21 presents descriptive statistics for students' pretest and posttest scores, showing a significant improvement in overall performance following the implementation of the ARGEONET medium. The mean score increased from 2.7 on the pretest to 3.633 on the posttest. These findings indicate an improvement in the indicator of students' inference on critical thinking. The standard deviation increased from 1.055364 to 1.473521, indicating a wider distribution of performance scores among participants following the use of ARGEONET. A one-group pretest-posttest design was used. First, a normality test was conducted on students' pretest and posttest responses for the inference indicator presented in Table 22.

**Table 22**. Test of normality

| Types of Test | Shapiro-Wilk | | |
|---|---|---|---|
| | Statistic | df | Sig. |
| Inference_Pretest | 0.660 | 30 | 0.000 |
| Inference_Posttest | 0.839 | 30 | 0.000 |

Table 22 shows that the results of the normality test using the Shapiro–Wilk test indicate a significance value (Sig.) of 0.000 for the inference indicator in both the pretest and posttest. Both significance values are less than the threshold of $\alpha = 0.05$ (Sig. $< 0.05$), indicating that the data in both groups are not normally distributed. Because the data are not normally distributed, a nonparametric test specifically, the Wilcoxon test. The results of the Wilcoxon test are presented in Table 23.

**Table 23**. Results of the Wilcoxon test

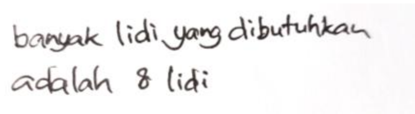
| | Inference_Posttest - Inference_Pretest |
|---|---|
| Z | -2.999 |
| Asymp. Sig. (2-tailed) | 0.003 |

Table 23 shows that the significance value (Sig.) of the Wilcoxon test is 0.003 for the inference indicator. This significance indicates that there are differences in students' critical thinking skills regarding the inference indicator. Furthermore, an effect size test was conducted to determine the extent to which the use of ARGEONET improved the achievement of the inference indicators. The effect size test revealed that $r = \frac{|Z|}{\sqrt{n}} = \frac{|-2.999|}{\sqrt{30}} = 0.655 > 0.50$. Therefore, the impact of using ARGEONET is large to improve the inference indicator of critical thinking.

Next, excerpts of pretest and posttest answers are presented, as well as excerpts from in-depth interviews with students who experienced an increase in scores on the inference indicator in critical thinking skills. One of the subjects who experienced an increase in score on the inference indicator, referred to by the researcher as S3. Figure 12 shows S3's answers in the pretest and posttest.


**Translation:**
Minimum wire length = 2592 $cm^2$

(a)




**Translation:**
The number of sticks needed is 8 sticks

(b)

**Figure 12**. (a) S3' pretest answer related to the inference indicator; (b) S3' posttest answers related to the inference indicator

Figure 12(a) shows S3's answer before the implementation of ARGEONET in learning. S3 wrote a conclusion with an incorrect final answer. On the other hand, Figure 12(b) shows S3's answers after using ARGEONET. S3 wrote the conclusion correctly. A more detailed explanation regarding the student's answers is presented in the excerpt of Interview 3.

*Interview 3*

*R : Why did you answer like that in concluding your answers on the pretest and posttest questions?*

*S3 : Because I adjust my calculation results*

Based on interview 3, S3 explained that the difference in conclusions between the pretest and posttest occurred because S3 adjusted them according to the calculation results before determining the final answer. After using ARGEONET, S3 was able to perform calculations accurately and make appropriate conclusions. This indicates an improvement in the student's reasoning ability, particularly in evaluating the accuracy of calculation results and drawing conclusions.

***Situation***

This section presents the results of a test of ARGEONET's effectiveness in improving the situation indicator. The first step in this analysis is to conduct a descriptive statistical analysis to determine the distribution of scores and assess student performance before and after the intervention. Table 24 presents descriptive statistics for students' pretest and posttest scores in the situation indicator.

**Table 24**. Descriptive statistics of pretest and posttest scores in the situation indicator

| Types of Tests | | Statistic | Std. Error |
|---|---|---|---|
| Pretest_Situation | Mean | 4.1 | 0.259752 |
| | Std. Deviation | 1.422722 | |
| | Minimum | 2.00 | |
| | Maximum | 6.00 | |
| Posttest_Situation | Mean | 4.6 | 0.247284 |
| | Std. Deviation | 1.354431 | |
| | Minimum | 1.00 | |
| | Maximum | 6.00 | |

Table 24 presents descriptive statistics for students' pretest and posttest scores, showing a significant improvement in overall performance following the implementation of the ARGEONET medium. The mean score increased from 4.1 on the pretest to 4.6 on the posttest. These findings indicate an improvement in the indicator of students' situation on critical thinking. The standard deviation decreased from 1.422722 to 1.354431, indicating reduced score variability and a more consistent level of achievement among participants following the use of ARGEONET. A one-group pretest-posttest design was used. First, a normality test was conducted on students' pretest and posttest responses for the situation indicator presented in Table 25.

**Table 25**. Test of normality

| Types of Test | Shapiro-Wilk | | |
|---|---|---|---|
| | Statistic | df | Sig. |
| Situation_Pretest | 0.847 | 30 | 0.001 |
| Situation_Posttest | 0.855 | 30 | 0.001 |

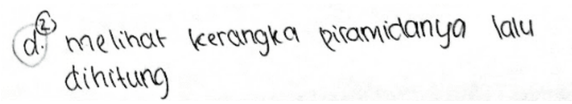
Table 25 shows that the results of the normality test using the Shapiro–Wilk test indicate a significance value (Sig.) of 0.001 for the situation indicator in both the pretest and posttest. Both significance values are less than the threshold of $\alpha = 0.05$ (Sig. $< 0.05$), indicating that the data in both groups are not normally distributed. Because the data are not normally distributed, a nonparametric test specifically, the Wilcoxon test. The results of the Wilcoxon test are presented in Table 26.

**Table 26**. Results of the Wilcoxon test

| | Situation_Posttest - Situation_Pretest |
|---|---|
| Z | -1.961 |
| Asymp. Sig. (2-tailed) | 0.050 |

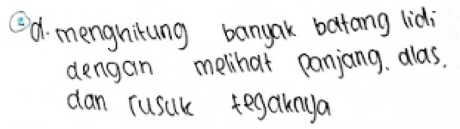
Table 26 shows that the significance value (Sig.) of the Wilcoxon test is 0.050 for the situation indicator. This significance indicates that there are differences in students' critical thinking skills regarding the situation indicator. Furthermore, an effect size test was conducted to determine the extent to which the use of ARGEONET improved the achievement of the situation indicators. The effect size test revealed that $r = \frac{|Z|}{\sqrt{n}} = \frac{|-1.961|}{\sqrt{30}} = 0.358 > 0.30$. Therefore, the impact of using ARGEONET is medium to improve the situation indicator of critical thinking.

Next, excerpts of pretest and posttest answers are presented, as well as excerpts from in-depth interviews with students who experienced an increase in scores on the situation indicator in critical thinking skills. One of the subjects who experienced an increase in score on the situation indicator, referred to by the researcher as S4. Figure 13 shows S4's answers in the pretest and posttest.




**Translation:**
Looking at the pyramid framework and then counted

(a)




**Translation:**
Calculating the number of sticks by looking at the length of the base and its vertical edges

(b)

**Figure 13**. (a) S4' pretest answer related to the situation indicator; (b) S4' posttest answers related to the situation indicator

Figure 13(a) shows S4's answer before the implementation of ARGEONET in learning. S4 uses some of the information that is relevant to the problem, but there is still information that is not utilized. The information that has not been included is the use of the pyramid edge length information. On the other hand, Figure 13(b) shows S4's answers after using ARGEONET. S4 uses some of the information that is relevant to the problem, namely information regarding the base edge length and the pyramid's vertical edge length. A more detailed explanation regarding the student's answers is presented in the excerpt of Interview 4.

*Interview 4*

*R : Why is your problem-solving approach different between the pretest and posttest?*

*S4 : In the pretest, I only used the information that I considered most important, but did not pay attention to all the information that I actually used to answer the questions. In the posttest, I better understand the shape of the pyramid, so I can use the information on the base edge and the vertical edge lengths needed to solve the problem.*

Based on interview 4, S4 indicated that using ARGEONET helped S4 in utilizing relevant information to answer questions, allowing S4 to use the necessary data more accurately in the problem-solving process. This supports that S4 experienced an improvement in critical thinking ability on the situation indicator.

***Clarity***

This section presents the results of a test of ARGEONET's effectiveness in improving the clarity indicator. The first step in this analysis is to conduct a descriptive statistical analysis to determine the distribution of scores and assess student performance before and after the intervention. Table 27 presents descriptive statistics for students' pretest and posttest scores in the clarity indicator.

**Table 27**. Descriptive statistics of pretest and posttest scores in the clarity indicator

| Types of Tests | | Statistic | Std. Error |
|---|---|---|---|
| Pretest_Clarity | Mean | 2.0 | 0.166091 |
| | Std. Deviation | 0.909718 | |
| | Minimum | 0.00 | |
| | Maximum | 5.00 | |
| Posttest_Clarity | Mean | 2.133 | 0.223521 |
| | Std. Deviation | 1.224276 | |
| | Minimum | 0.00 | |
| | Maximum | 4.00 | |

Table 27 presents descriptive statistics for students' pretest and posttest scores, showing a significant improvement in overall performance following the implementation of the ARGEONET medium. The mean score increased from 2.0 on the pretest to 2.133 on the posttest. These findings indicate an improvement in the indicator of students' clarity on critical thinking. The standard deviation increased from 0.909718 to 1.224276, indicating a wider distribution of performance scores among participants following the use of ARGEONET. A one-group pretest-posttest design was used. First, a normality test was conducted on students' pretest and posttest responses for the clarity indicator presented in Table 28.

**Table 28**. Test of normality

| Types of Test | Shapiro-Wilk | | |
|---|---|---|---|
| | Statistic | df | Sig. |
| Clarity_Pretest | 0.847 | 30 | 0.000 |
| Clarity_Posttest | 0.855 | 30 | 0.022 |

Table 28 shows that the results of the normality test using the Shapiro–Wilk test indicate a significance value (Sig.) of 0.000 for the clarity indicator in the pretest and a significance value (Sig.) of 0.022 for the clarity indicator in the posttest. Both significance values are less than the threshold of $\alpha = 0.05$ (Sig. < 0.05), indicating that the data in both groups are not normally distributed. Because the data are not normally distributed, a nonparametric test specifically, the Wilcoxon test. The results of the Wilcoxon test are presented in Table 29.

**Table 29**. Results of the Wilcoxon test

| | Clarity_Posttest - Clarity_Pretest |
|---|---|
| Z | -0.705 |
| Asymp. Sig. (2-tailed) | 0.481 |

Table 29 shows that the significance value (Sig.) of the Wilcoxon test is 0.481 for the clarity indicator. These non-significant results indicate that there is no difference in students' critical thinking skills regarding the clarity indicator. Furthermore, an effect size test was conducted to determine the extent to which the use of ARGEONET improved the achievement of the clarity indicators. The effect size test revealed that $r = \frac{|Z|}{\sqrt{n}} = \frac{|-0.705|}{\sqrt{30}} = 0.129 > 0.10$. Therefore, the impact of using ARGEONET is small to improve the situation indicator of critical thinking.

Next, excerpts of pretest and posttest answers are presented, as well as excerpts from in-depth interviews with students who experienced an increase in scores on the clarity indicator in critical thinking skills. One of the subjects that showed changes in clarity indicators, referred to by the researcher as S5. Figure 14 shows S5's answers in the pretest and posttest.

| (a) | (b) |
|---|---|
| Karena untuk mencari kerangka yang dibutuhkan hanya panjang rusuk & rusuk tegak. | Karena kerangka limas segi 4, 8 yang tegak dari 4 persegi 4 sisi tegak untuk mencari kerangka yang dibutuhkan panjang keliling |
| **Translation:** Because to find the required frame, only the length of the sides and the vertical sides are needed | **Translation:** Because the square pyramid frame, 8 consisting of 4 squares and 4 upright sides, to find the frame the required perimeter length |

Figure 14. (a) S5' pretest answer related to the clarity indicator (b) S5' posttest answers related to the clarity indicator

Figure 14(a) shows S5's answer before the implementation of ARGEONET in learning. S5 explained that to find the frame, the length of the sides and the length of the vertical sides are needed. However, S5 has not explained the relationship between this information and the goal that is to be achieved. On the other hand, Figure 14(b) shows S5's answers after using ARGEONET. S5 has begun to show improvement by identifying the pyramid's edge parts to support the conclusion he made. Nevertheless, the given answer is still incorrect, because the base of a square pyramid consists of only one side, which is a square, and four vertical sides that are triangles. A more detailed explanation regarding the student's answers is presented in the excerpt of Interview 5.

*Interview 5*

*R* : *Why did you use the side length and the slant height to determine the pyramid's frame? What is the relationship with the answer you are looking for?*

*S5* : *Because to find the pyramid's frame, what is needed is the length of all its edges. So, I used the base side length and the slant length to calculate the entire frame.*

Based on the answers from S5 and Interview 5, although S5 showed improvement in providing more structured explanations after using ARGEONET, the reasons given were still not mathematically accurate. This indicates that S5 has made progress in critical thinking abilities on the clarity indicator, where S5 has begun to be able to direct relevant information to support conclusions, but still needs reinforcement in understanding the characteristics of three-dimensional shapes.

***Overview***

This section presents the results of a test of ARGEONET's effectiveness in improving the overview indicator. The first step in this analysis is to conduct a descriptive statistical analysis to determine the distribution of scores and assess student performance before and after the intervention. Table 30 presents descriptive statistics for students' pretest and posttest scores in the overview indicator.

**Table 30**. Descriptive statistics of pretest and posttest scores in the overview indicator

| Types of Tests | | Statistic | Std. Error |
|---|---|---|---|
| Pretest_Overview | Mean | 2.066667 | 0.262321 |
| | Std. Deviation | 1.436791 | |
| | Minimum | 0.00 | |
| | Maximum | 5.00 | |
| Posttest_Overview | Mean | 2.7 | 0.34624 |
| | Std. Deviation | 1.896458 | |
| | Minimum | 0.00 | |
| | Maximum | 6.00 | |

Table 30 presents descriptive statistics for students' pretest and posttest scores, showing a significant improvement in overall performance following the implementation of the ARGEONET medium. The mean score increased from 2.067 on the pretest to 2.7 on the posttest. These findings indicate an improvement in the indicator of students' overview on critical thinking. The standard deviation increased from 1.436791 to 1.896458, indicating a wider distribution of performance scores among participants following the use of ARGEONET. A one group pretest-posttest design was used. First, a normality test was conducted on students' pretest and posttest responses for the overview indicator presented in Table 31.

**Table 31**. Test of normality

| Types of Test | Shapiro-Wilk | | |
|---|---|---|---|
| | Statistic | df | Sig. |
| Overview_Pretest | 0.899 | 30 | 0.008 |
| Overview_Posttest | 0.890 | 30 | 0.005 |

Table 31 shows that the results of the normality test using the Shapiro–Wilk test indicate a significance value (Sig.) of 0.008 for the overview indicator in the pretest and a significance value (Sig.) of 0.005 for the overview indicator in the posttest. Both significance values are less than the threshold of $\alpha = 0.05$ (Sig. < 0.05), indicating that the data in both groups are not normally distributed. Because the data are not normally distributed, a nonparametric test specifically, the Wilcoxon test. The results of the Wilcoxon test are presented in Table 32.

**Table 32**. Results of the Wilcoxon test

| | Overview_Posttest - Overview_Pretest |
|---|---|
| Z | -2.033 |
| Asymp. Sig. (2-tailed) | 0.042 |

Table 32 shows that the significance value (Sig.) of the Wilcoxon test is 0.042 for the overview indicator. This significance indicates that there are differences in students' critical thinking skills regarding overview indicator. Furthermore, an effect size test was conducted to determine the extent to which the use of ARGEONET improved the achievement of the overview indicators. The effect size test revealed that $r = \frac{|Z|}{\sqrt{n}} = \frac{|-2.033|}{\sqrt{30}} = 0.371 > 0.30$. Therefore, the impact of using ARGEONET is medium to improve the overview indicator of critical thinking.

Next, excerpts of pretest and posttest answers are presented, as well as excerpts from in-depth interviews with students who experienced an increase in scores on the overview indicator in critical thinking skills. One of the subjects who experienced an increase in score on the overview indicator, referred to by the researcher as S6. Figure 15 shows S6's answers in the pretest and posttest.


**Translation:**
I am already confident with that answer

(a)

**Translation:**
Redo it more carefully and if the answer is the same then the answer is not wrong

(b)

**Figure 15**. (a) S6' pretest answer related to the overview indicator (b) S5' posttest answers related to the overview indicator

Figure 15(a) shows S6's answer before the implementation of ARGEONET in learning. S6 explained that without doing a check, S6 was already confident in the answer S6 wrote. On the other hand, Figure 15(b) shows S6's answers after using ARGEONET. S6 did a check by redoing the problem and ensuring that the initial and final answers were the same. A more detailed explanation regarding the student's answers is presented in the excerpt of Interview 6.

*Interview 6*

*R* : *In the pretest, you wrote that you did not do a check after finding the final answer. What made you confident not to do a check?*

*S6* : *Because I feel that the steps I took are already in accordance with the way I understand, I think the answer is correct and does not need to be recalculated.*

*R* : *So, why did you perform the check during the posttest?*

*S6* : *Because I am more aware that the answers need to be double-checked. Like when solving problems in its medium. I redo it in the same way, then compare the results with the first answer.*

Based on interview 6, S6's answer is reinforced by the statement that after using ARGEONET, awareness of checking increased because there is a feature in ARGEONET that habituates that activity. Thus, S6 shows an improvement in critical thinking skills on the overview indicator.

### 3.2. Discussion

The product of this research and development (R&D) is a learning media named ARGEONET, designed to enhance students' critical thinking. The development process involved a comprehensive literature review, analysis of comparable learning media, and students interviews to identify user needs. To ensure alignment with the research objectives, ARGEONET underwent expert validation. The validation results from two experts yielded average scores of 4.45 (89%) and 4.65 (93%), which are classified as very valid. These findings demonstrate that ARGEONET was developed with e-content and media features well-suited to students' needs and met high-quality standards in terms of both content relevance and media feasibility (Lin & Yu, 2023).

Next, the practicality test of the media was conducted to test the quality of ARGEONET based on ease of use, feasibility, and the benefits of the features available in supporting the learning process. The results of the practicality assessment revealed that the ARGEONET media achieved a score percentage of 86.06% based on students' responses and 91.43% from teachers. These outcomes suggest that the developed media is categorized as practical for classroom implementation. This finding aligns with previous research indicating that learning media are considered practical when their usability level exceeds the minimum threshold of 61%, demonstrating their effectiveness for instructional use (Riduwan, 2022). Moreover, the high score in design attractiveness and interactivity (4.24), along with material relevance (3.00), indicate that ARGEONET is able to meet students' needs by providing visually engaging and interactive features, as well as content that aligns with the intended learning objectives. This is consistent with previous studies indicating that interactive and student-relevant learning media can enhance students' learning motivation, thereby fostering the development of their critical thinking skills (Al Hajri et al., 2025; Djamas et al., 2018).

The effectiveness of ARGEONET shows an improvement in students' critical thinking skills on all indicators, except the clarity indicator. Of the six indicators available, only 3 indicators experienced an increase with a large effect size, namely the indicators focus, reason, and inference. Meanwhile, 2 indicators experienced an increase with a medium effect size (situation and overview) and 1 indicator did not experience an increase (clarity). This shows that ARGEONET has a different influence on each indicator. This difference is influenced by the features within ARGEONET that facilitate certain indicators, so indicators that have greater support in terms of activity in its media will produce a larger effect size compared to other indicators (Lin & Yu, 2023).

Effectiveness tests showed that there was a significant difference in the focus indicator before and after using ARGEONET with the large effect size. This improvement is supported by the ARGEONET menu, specifically in Activity 1 and Activity 2. This activity also facilitates the first M6 learning model syntax. This activity focuses on focus indicator and initial abilities since students carry out the identification of relevant information by understanding problems and the basic concepts of nets of three-dimensional shapes, namely the elements of cubes, blocks, prisms, and pyramids. The interactive videos in activity 1 provide opportunities for students to explore and understand concepts with flexible access to learning resources, whereas the augmented reality feature enhances conceptual visualization. Visualization through augmented reality facilitates students in observing objects presented from various directions or perspectives. This can minimize errors in understanding geometric concepts and help form a deep comprehension (Rohendi & Wihardi, 2020; Tarng et al., 2024).

Effectiveness tests showed that there was a significant difference in the reason indicator before and after using ARGEONET with the large effect size. This improvement is supported by the ARGEONET menu, specifically in Activity 3. Activity 3 also facilitates the third syntax in the M6 learning model. In this part, students engage in discussions to explain the strategies and steps used to solve the problems presented. This section also provides AR content that offers simulations of three-dimensional shapes transforming into nets. Through these simulations, students can observe the relationships between the elements of the three-dimensional shapes and their nets, which helps them construct logical explanations based on the observed processes (Chen et al., 2024; Molle & Sabandar, 2025). In addition, the discussion activities encourage students to express relevant strategies and reasoning when answering the given problems. Thus, the integration of AR visualization and discussion activities in ARGEONET contributes to strengthening students' ability to provide reasoning when solving geometry problems.

Effectiveness tests showed that there was a significant difference in the inference indicator before and after using ARGEONET with the large effect size. This improvement is supported by the ARGEONET menu, specifically in Activity 4 dan 6. This activity also facilitates the sixth syntax in the M6 learning model. In these sections, students make conclusions using their own language. Through this activity, students are encouraged to connect the results of their exploration, analysis, and previous reasoning processes to produce a logical conclusion. Therefore, if the previous process is wrong, it will affect the conclusion made.

Unlike the previous three indicators, although it also shows an increase before and after the use of ARGEONET, the situation indicator has a medium effect size. This improvement is supported by the ARGEONET menu, specifically in Activity 3. In this section, in addition to facilitating the reason indicator, it also facilitates the inference indicator because there are activities identifying the information needed to answer the characteristics of certain three-dimensional shapes. On the other hand, the clarity indicator did not experience a significant change before and after using ARGEONET. The effect size it had on using ARGEONET was also small, so it did not have a major influence on improving critical thinking skills for the clarity indicator. When linked to the menus available on ARGEONET, this indicator is not explicitly facilitated. Support for the clarity indicator on ARGEONET is implied in Activity 3. This is seen when students provide reasons for choosing strategies or steps used in solving problems. However, this support has not been specifically designed to support the clarity indicator, so students' improvement in critical thinking for the clarity indicator has not shown a significant increase.

Like the situation indicator, effectiveness tests showed that there was a significant difference in the overview indicator before and after using ARGEONET with the medium effect size. This improvement is supported by the ARGEONET menu, specifically in Activity 5. This activity also facilitates the fifth syntax in the M6 learning model. In this section, students carry out evaluations by identifying which answers they consider wrong and which answers they consider correct. This can be done by students if they conduct a thorough check. Such activities encourage learners to monitor and evaluate their own thinking processes, which are important components of self-regulated learning and critical thinking (Dignath et al., 2023; Hyytinen et al., 2023). With this menu, students get used to performing checks and it can support their thinking skills on the overview indicator.

Based on the effectiveness test results for each indicator, there are 5 critical thinking indicators that show a significant difference before and after using ARGEONET, this indicates that the use of ARGEONET is effective for enhancing students' critical thinking skills. This finding is in line with O'Connor and Mahony (2023), who reported that the integration of augmented reality (AR) in learning environments enhances student engagement and significantly improves higher-order thinking skills, including critical thinking. Similarly, Chang et al. (2022) conducted a meta-analysis demonstrating that AR technologies have a consistently positive impact on students' learning performance and cognitive development. In addition, Huang (2024) found that AR-based simulation learning improves clinical readiness and critical thinking skills among nursing students, reinforcing the effectiveness of ARGEONET in promoting critical thinking. Therefore, these empirical findings confirm that augmented reality-based learning media can be considered an effective pedagogical approach for enhancing students' critical thinking skills across various educational contexts.

There are limitations in this study where the number of participants involved was limited to only one class without a comparison group, so the results obtained cannot be generalized to a wider population. Therefore, further research can be conducted by involving a comparison group and more participants, at least from one province, in order to obtain more comprehensive results.

## 4. CONCLUSION

This research concludes that ARGEONET was valid, practical, and effective for implementation in mathematics learning. The validation results from two experts yielded average scores of 4.45 (89%) and 4.65 (93%), which are classified as very valid. Besides, the results of the practicality assessment revealed that the ARGEONET media achieved a score percentage of 86.06% based on students' responses and 91.43% from teachers, which are classified as very practical. The effectiveness test of ARGEONET showed that out of six critical thinking indicators, five of them showed significant differences before and after using ARGEONET. The improvements were in the indicators of focus, reason, inference, situation, and overview. Meanwhile, the clarity indicator did not show a significant difference before and after using ARGEONET, which was caused by the lack of explicit support in ARGEONET. However, overall, ARGEONET successfully stimulated higher-order thinking through exploration, analysis, evaluation, and reflection of geometric concepts. Therefore, ARGEONET shows potential as an augmented reality–based pedagogical approach for improving students' critical thinking skills in geometry learning. However, this study has limitations in the scope of trials, which were restricted to junior high school students in Banyuwangi Regency, East Java, Indonesia. Furthermore, this study where the number of participants involved was limited to only one class without a comparison group, so the results obtained cannot be generalized to a wider population. Therefore, further research can be conducted by involving a comparison group and more participants, at least from one province, in order to obtain more comprehensive results.

## Acknowledgments

The authors would like to thank the Universitas Negeri Malang for funding under contract number 14.4.64/UN32.14.1/LT/2025. This work was supported by the Indonesian Collaborative Research Grant (RKI) Scheme A 2025, involving Universitas Negeri Malang, Universitas Terbuka, and Universitas Pendidikan Indonesia.

## Declarations

| | |
|---|---|
| Author Contribution | : PD: Conceptualization, Investigation, Methodology, and Writing - original draft; IR: Conceptualization, Data curation, Project administration, and Writing - review & editing; TS: Supervision, and Validation; ADS: Formal analysis, and Visualization; TK: Resources, and Writing - review & editing; SF: Formal analysis, and Validation; EP: Supervision, and Validation. |
| Funding Statement | : This research was funded by Universitas Negeri Malang through the Indonesian Collaborative Research Grant (RKI) Scheme A 2025 with contract number 14.4.64/UN32.14.1/LT/2025. |
| Conflict of Interest | : The authors declare no conflict of interest. |
| Additional Information | : Additional information is available for this paper. |

## REFERENCES

Al Hajri, M., Bani Orabah, S., Domalaon, D., & Bijani, H. (2025). Evaluating the quality of blended learning in higher education: Teachers' and students' perceptions in an Omani EFL context. *Arab World English Journal*, *11*, 20–42. https://doi.org/10.24093/awej/call11.2

Al Khateeb, A., & Alotaibi, H. (2024). The upsurge of the metaverse in educational settings: a meta-analysis study. *Pegem Journal of Education and Instruction*, *14*(3), 62–75.

Alruthaya, A., Nguyen, T.-T., & Lokuge, S. (2021). The application of digital technology and the learning characteristics of Generation Z in higher education *Australasian Conference on Information Systems*, Sidney.

Aslan, S., & Aybek, B. (2024). Development of critical-thinking skills rubric within the scope of multicultural education. *Educational Process International Journal*, *13*(3), 139–158. https://doi.org/10.22521/edupij.2024.133.8

Avci, Ü., & AkgÜL, F. (2024). The effect of metaverse technologies on education and human-computer interaction: A systematic analysis of the literature. *Journal of Educational Technology and Online Learning*, *7*(1), 1–16. https://doi.org/10.31681/jetol.1302048

Avcu, Y. E., & Yaman, Y. (2024). Differentiated science instruction in the metaverse with gifted students: Let's visit Kaz Mountains national park! *Journal of Inquiry Based Activities*, *14*(2), 46–65.

Chang, H.-Y., Binali, T., Liang, J.-C., Chiou, G.-L., Cheng, K.-H., Lee, S. W.-Y., & Tsai, C.-C. (2022). Ten years of augmented reality in education: A meta-analysis of (quasi-) experimental studies to investigate the impact. *Computers & Education*, *191*, 104641. https://doi.org/10.1016/j.compedu.2022.104641

Cheerapakorn, P., Hinon, K., & Wannapiroon, P. (2024). Hybrid project-based learning model on metaverse to enhance collaboration. *International Education Studies*, *17*(6), 65–78. https://doi.org/10.5539/ies.v17n6p65

Chen, S.-J., Chen, C.-Q., & Shan, X.-F. (2024). The effects of an immersive virtual-reality-based 3D modeling approach on the creativity and problem-solving tendency of elementary school students. *Sustainability*, *16*(10), 4092. https://doi.org/10.3390/su16104092

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Demircioglu, T., Karakus, M., & Ucar, S. (2023). Developing students' critical thinking skills and argumentation abilities through augmented reality–based argumentation activities in science classes. *Science & Education*, *32*(4), 1165–1195. https://doi.org/10.1007/s11191-022-00369-5

Dignath, C., van Ewijk, R., Perels, F., & Fabriz, S. (2023). Let learners monitor the learning content and their learning behavior! A meta-analysis on the effectiveness of tools to foster monitoring. *Educational Psychology Review*, *35*(2), 62. https://doi.org/10.1007/s10648-023-09718-4

Djamas, D., Tinedi, V., & Yohandri, Y. (2018). Development of interactive multimedia learning materials for improving critical thinking skills. *International Journal of Information and Communication Technology Education*, *14*(4), 66–84. https://doi.org/10.4018/ijicte.2018100105

Egan, L., Tang, J. H., Ferraro, D., Erberber, E., Tsokodayi, Y., Stearns, P., & Malley, L. (2022). *US technical report and user guide for the 2019 trends in international mathematics and science study (TIMSS). Part 1. NCES 2022-049*. National Center for Education Statistics. https://files.eric.ed.gov/fulltext/ED623028.pdf

Ennis, R. H. (1996). *Critical thinking*. Prentice-Hall Inc.

Eşin, Ş., & Özdemir, E. (2022). The metaverse in mathematics education: The opinions of secondary school mathematics teachers. *Journal of Educational Technology and Online Learning*, *5*(4), 1041–1060. https://doi.org/10.31681/jetol.1149802

Facione, P. A., Sanchez, C. A., Facione, N. C., & Gainen, J. (1995). The disposition toward critical thinking. *The Journal of general education*, *44*(1), 1–25. https://www.jstor.org/stable/27797240

Groenwold, R. H. H., Shofty, I., Miočević, M., van Smeden, M., & Klugkist, I. (2018). Adjustment for unmeasured confounding through informative priors for the confounder-outcome relation. *BMC Medical Research Methodology*, *18*(1), 174. https://doi.org/10.1186/s12874-018-0634-3

Hazaymeh, W. A., & Alomery, M. K. (2021). The effectiveness of visual mind mapping strategy for improving english language learners' critical thinking skills and reading ability. *European Journal of Educational Research*, *11*(1), 141–150. https://doi.org/10.12973/eu-jer.11.1.141

Huang, H.-M. (2024). Effectiveness of simulation-based augmented reality in enhancing pediatric nursing and clinical reasoning competency among students: A quasi-experimental study. *Clinical Simulation in Nursing*, *95*, 101601. https://doi.org/10.1016/j.ecns.2024.101601

Hyytinen, H., Nissinen, K., Kleemola, K., Ursin, J., & Toom, A. (2023). How do self-regulation and effort in test-taking contribute to undergraduate students' critical thinking performance? *Studies in Higher Education*, *49*(1), 192–205. https://doi.org/10.1080/03075079.2023.2227207

Kemendikbud. (2022). *Panduan pembelajaran dan asesmen di sekolah menengah pertama [Learning and assessment guide in junior high school]*. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.

Khusna, A. H., Siswono, T. Y. E., & Wijayanti, P. (2024). Mathematical problem design to explore students' critical thinking skills in collaborative problem solving. *Mathematics Teaching Research Journal*, *16*(3), 217–240.

Kishore, K., Jaswal, V., Kulkarni, V., & De, D. (2021). Practical guidelines to develop and evaluate a questionnaire. *Indian Dermatology Online Journal*, *12*(2), 266–275. https://doi.org/10.4103/idoj.IDOJ_674_20

Kongkhen, P., & Chatwattana, P. (2023). The project-based learning: PjBL via brainstorming with metaverse to promote multimedia production skills. *Higher Education Studies*, *13*(4), 193–200. https://doi.org/10.5539/hes.v13n4p193

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lin, Y., & Yu, Z. (2023). Extending technology acceptance model to higher-education students' use of digital academic reading tools on computers. *International Journal of Educational Technology in Higher Education*, *20*(1), 34. https://doi.org/10.1186/s41239-023-00403-8

Ma, J., Jin, Y., Zhao, D., & Chen, L. (2023). The research on the application of educational metaverse in the experimental teaching of medical physiology. *International Journal of Technology in Teaching and Learning*, *19*(1), 69–85. https://doi.org/10.37120/ijttl.2023.19.1.04

Mahdi, H. S., & Sahari, Y. M. (2024). Exploring the relationship between critical thinking, attitude, and anxiety in shaping the adoption of artificial intelligence in translation among Saudi translators. *Journal of Pedagogical Research*, *8*(2), 81–94. https://doi.org/10.33902/jpr.202423345

Mandala, A. S., Anwar, L., Sa'dijah, C., & Zulnaidi, H. (2025). Development of mobile augmented reality-based geometry learning games to facilitate spatial reasoning. *Infinity Journal*, *14*(2), 323–348. https://doi.org/10.22460/infinity.v14i2.p323-348

Maulida, A. S., Wahyudin, W., Turmudi, T., & Nurlaelah, E. (2024). The effect of experiential learning and directed instructions assisted by augmented reality on students' self-regulated learning. *Infinity Journal*, *13*(2), 553–568. https://doi.org/10.22460/infinity.v13i2.p553-568

Molle, J. S., & Sabandar, V. P. (2025). Developing the reasoning ability of students through three-dimensional figures learning *The 2nd International Conference on Mathematics and Mathematics Education (ICMMeD) 2023*,

Mutchima, P., Leelakitpaisarn, Y., Pijitkamnerd, B., Phiwma, N., & Pantrakool, S. (2025). The effect of metaverse technology on multicultural learning: Strengthening the social attitudes, cultural awareness and critical thinking skills of secondary school students. *Journal of Education and e-Learning Research*, *12*(1), 21–30. https://doi.org/10.20448/jeelr.v12i1.6329

Nieveen, N., & Folmer, E. (2013). Formative Evaluation in Educational Design Research. In T. Plomp & N. Nieveen (Eds.), *Educational design research* (pp. 152–169). Netherlands Institute for Curriculum Development (SLO).

Nosrat, A., & Parsafar, A. (2018). Effect of problem-solving based methods on critical thinking and academic achievement of math among students of saravan higher education centers. *International Journal of Applied Mathematical Research*, *7*(1), 1–6. https://doi.org/10.14419/ijamr.v7i1.8397

O'Connor, Y., & Mahony, C. (2023). Exploring the impact of augmented reality on student academic self-efficacy in higher education. *Computers in human Behavior*, *149*, 107963. https://doi.org/10.1016/j.chb.2023.107963

OECD. (2023). *PISA 2022 results (volume I): The state of learning and equity in education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Pereles, A., Ortega-Ruipérez, B., & Lázaro, M. (2024). The power of metacognitive strategies to enhance critical thinking in online learning. *Journal of Technology and Science Education*, *14*(3), 831–843. https://doi.org/10.3926/jotse.2721

Plomp, T. (2013). Educational design research: An introduction. In T. Plomp & N. Nieveen (Eds.), *Educational design research* (pp. 10–51). Netherlands Institute for Curriculum Development (SLO).

Rachmadtullah, R., Setiawan, B., Wasesa, A. J. A., Wicaksono, J. W., & Rasmitadila, R. (2023). The utilization of metaverse technology applications based on science, technology, engineering and mathematics (Meta-STEM) to improve critical thinking skills. *Journal of Education and e-Learning Research*, *10*(4), 778–784. https://doi.org/10.20448/jeelr.v10i4.5203

Rassyi, S. F. R., Supriadi, S., Andayani, Y., Hakim, A., Burhanuddin, B., & Hadisaputra, S. (2023). Development of the interactive learning media based on augmented reality 3D on the petroleum concept. *IJCER (International Journal of Chemistry Education Research)*, *7*(1), 44–51. https://doi.org/10.20885/ijcer.vol7.iss1.art8

Riduwan, R. (2022). *Skala pengukuran variabel-variabel penelitian [Measurement scale of research variables]*. Alfabeta.

Rohendi, D., & Wihardi, Y. (2020). Learning three-dimensional shapes in geometry using mobile-based augmented reality. *International Journal of Interactive Mobile Technologies (iJIM)*, *14*(9), 48–60. https://doi.org/10.3991/ijim.v14i09.13035

Suh, W., & Ahn, S. (2022). Utilizing the metaverse for learner-centered constructivist education in the post-pandemic era: An analysis of elementary school students. *Journal of Intelligence*, *10*(1), 17. https://doi.org/10.3390/jintelligence10010017

Susandi, A. D., Sa'dijah, C., As'ari, A. R., & Susiswo, S. (2022). Developing the M6 learning model to improve mathematic critical thinking skills. *Pedagogika*, *145*(1), 182–204. https://doi.org/10.15823/p.2022.145.11

Talan, T., & Kalinkara, Y. (2022). Students' opinions about the educational use of the metaverse. *International Journal of Technology in Education and Science*, *6*(2), 333–346. https://doi.org/10.46328/ijtes.385

Tarng, W., Huang, J.-K., & Ou, K.-L. (2024). Improving elementary students' geometric understanding through augmented reality and its performance evaluation. *Systems*, *12*(11), 493. https://doi.org/10.3390/systems12110493

Tural, A., & Koçak, N. (2023). Awareness levels of social studies pre-service teachers regarding metaverse use. *Advanced Education*, *11*(23), 69–86. https://doi.org/10.20535/2410-8286.284683

Wang, Q., & Abdullah, A. H. (2024). Enhancing students' critical thinking through mathematics in higher education: A systemic review. *SAGE Open*, *14*(3), 1–15. https://doi.org/10.1177/21582440241275651

Whitney-Smith, R. M. (2023). The emergence of computational thinking in national mathematics curricula: An Australian example. *Journal of Pedagogical Research*, *7*(2), 41–55. https://doi.org/10.33902/jpr.202318520

Wijaya, T. T., Hidayat, W., Hermita, N., Alim, J. A., & Talib, C. A. (2024). Exploring contributing factors to PISA 2022 mathematics achievement: Insights from Indonesian teachers. *Infinity Journal*, *13*(1), 139–156. https://doi.org/10.22460/infinity.v13i1.p139-156