Volume 4, Issue 3 (2026) 425-436

EQUATOR SCIENCE JOURNAL (ESJ)
e-ISSN: 2988-3881
Journal homepage:
https://ejournal.edukhatulistiwa.com/index.php/ESJ

Development of Problem-Solving Skills Instrument in the Context of Toroan Waterfall: A Validity Study

Lailatul Qodriyah[1], Herlina Mulyastuti[1*], Wiwin Puspita Hadi[1]
[1]Science Education Study Program, Universitas Trunodjoyo Madura, Indonesia

Received: September 7, 2026 | Revised: October 14, 2026 | Accepted: September 29, 2026
First Available Online: October 31, 2026 | Publication Date: October 31, 2026
DOI: https://doi.org/10.61142/esj.v4i3.496

*Corresponding author: herlina.mulyastuti@trunojoyo.ac.id

Keywords:
Problem-solving skills;
Validity;
Test instrument;
Toroan waterfall

ABSTRACT

This study aims to develop a problem-solving skills test instrument based on the local context of Toroan Waterfall and examine its content validity and students' responses to the developed instrument. The study employed a Research and Development (R&D) method using the ADDIE model, which consists of five stages: analysis, design, development, implementation, and evaluation. However, the implementation stage was not conducted because the study focused on developing and assessing the initial feasibility of the instrument. The developed instrument consisted of 10 essay questions covering work, energy, and simple machines, structured according to Polya's four problem-solving stages: understanding the problem, planning a strategy, implementing the plan, and reviewing the results. Content validity was assessed by four experts using Aiken's V formula, while students' responses were collected through one-to-one testing involving three students and small-group testing involving ten students. The results showed that the instrument achieved an average Aiken's V coefficient of 0.96, indicating a high level of content validity. Furthermore, students' response percentages reached 72% in the one-to-one test and 73% in the small-group test, both categorized as good. These findings indicate that the developed instrument has high content validity and received positive responses from students. Therefore, it is considered suitable for initial use as an assessment instrument for students' problem-solving skills in science learning, particularly in the context of Toroan Waterfall. Further research is needed to implement the instrument in classroom settings and examine its broader applicability.

This is an open access article under the CC BY-SA license.

INTRODUCTION

The 21st-century education requires students to have a myriad of competencies, one of which is problem-solving skills (Sari & Tanjung, 2022; Darmawati et al., 2025). This is particularly essential for navigating the intricacies of the present world and is fundamental to the development of a proficient generation (Amiruddin et al., 2024). The Merdeka Curriculum is the first innovative educational policy tailored for the 21st century, helping students pursue and develop their potential autonomously while strengthening competencies such as literacy and numeracy, as well as character education (Muizz et al., 2023).

Learners need to develop problem-solving skills, as it is the bedrock of scientific reasoning. This distinguishes one's ability to utilize knowledge and skills to identify, analysis, and develop answers for intricate issues (Polya, 1957; Siswanto & Meilisari, 2024). Problem-solving as a skill involves identifying the problem, devising a solution, implementing the solution, and assessing the solution's effectiveness (Polya, 1957; Utami & Puspitasari, 2022). Based on the 2022 PISA (Programme for International Student Assessment) survey, the unsatisfactory average of Indonesian students' problem-solving skills placed the country 67th out of 81 countries (OECD, 2023). Although there was a nominal increase in Indonesia's rank from PISA 2018, the country was still considered one of the bottom 15 countries globally (OECD, 2023). This justifies the need to prioritize strengthening problem-solving skill in science education in Indonesia.

Problem-solving skills are an important aspect that should be the focus of science learning; however, their application in schools remains limited. Based on research by Krismawati et al (2023), it was found that the implementation of problem-solving skills in schools, especially in the science learning process, has not been optimal. The primary factor contributing to this is the orientation of science learning, which often employs memorization-based methods and fails to emphasize the development of students' scientific process skills. Daily learning exercises have not been directed toward contextual problem-solving. Interviews with science teachers at a junior high school in Sampang Regency revealed that the integration of local wisdom contexts relevant to students' real-life experiences has never been applied in science learning, and students have not received specific exercises aimed at developing problem-solving skills. In addition, based on a questionnaire distributed, the majority of students stated that their problem-solving skills had never been specifically measured.

Evaluating a student’s problem-solving skills is one of the constituents of learning. This process provides information about students’ level of comprehension of the scientific knowledge and their application of the knowledge in their daily lives. The data collected from these measurements could serve to enhance the learning process and its assessment (Nurfadillah et al., 2023). The process of assessing problem-solving skills can utilize tools. Tools are designed to serve as evaluative aids in guiding students to achieve their learning goals more effectively, considering the context and circumstances of each individual (Chasanah et al., 2022). Such tools could take the form of science questions or problems that students are required to analysis and solve (Putra et al., 2025). The use of questions that resolve a problem and are rooted in the local wisdom gives these tools contextual and enhanced significance (Nirmalasari et al., 2022). Several prior studies also provide information indicating that, prior to formal education, the surrounding cultural environment serves as the first teacher to students (Murti & Sunarti, 2021). Therefore, it is possible to assist students in mastering the subject content to be taught by integrating local wisdom with the learning tools and their implementation. potential and kinetic energy of falling water, the work done by water, or the principles of simple machines that may be applied in the infrastructure around the waterfall. This material is part of the 8th-grade science curriculum in the odd semester.

Previous studies have explored the development of problem-solving assessment instruments by integrating local cultural contexts. Developed a problem-solving skills instrument on the topic of solid geometry by incorporating elements of the Baduy traditional house as a learning context (Khaerunnisa et al., 2025). Similarly, Tanjung et al. (2024) developed problem-solving test instruments based on cultural values in mechanical materials. These studies demonstrate the potential of integrating local contexts into assessment instruments. However, the use of natural phenomena at Toroan Waterfall as a context for developing a science problem-solving skills instrument covering work, energy, and simple machines has not been addressed in the studies discussed above. Therefore, this study focuses on

developing an instrument that connects the physical phenomena of Toroan Waterfall with the assessment of students' problem-solving skills.

The purpose of this study is to design an assessment tool in the form of a problem-solving skills test instrument using the Toroan Waterfall for Grade VIII science learning, particularly in the topics of work, energy, and simple machines. The study also examines the content validity of the developed instrument and students' responses through individual and small-group trials. The instrument is structured according to Polya's problem-solving stages: understanding the problem, devising a plan, carrying out the plan, and looking back. This study is expected to contribute an alternative contextual assessment instrument that connects local environmental phenomena with science concepts and provides an initial basis for assessing students' problem-solving skills.

METHOD

This study employed a research and development (R&D) approach to develop a problem-solving skills assessment instrument for Grade VIII science learning. The development process was guided by the ADDIE model, which consists of five stages: analysis, design, development, implementation, and evaluation (Herawati et al., 2024). Figure 1 illustrates the ADDIE development model, which served as the framework for this study, although only the analysis, design, and development stages were implemented, followed by preliminary evaluation through expert validation and student response trials. The implementation stage was not conducted because the study focused on developing the instrument and examining its initial feasibility rather than testing its effectiveness in classroom learning.

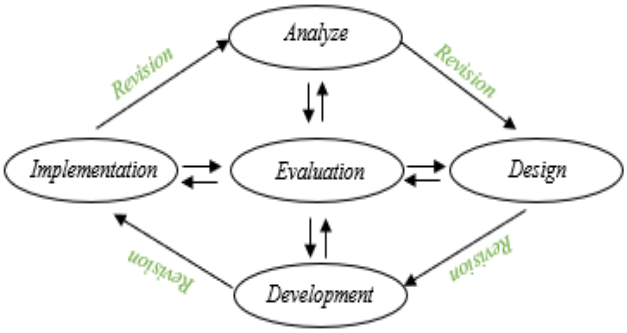


Fig.1. ADDIE Development Model

The analysis stage involved identifying the learning potential of Toroan Waterfall, student characteristics, and learning needs. These data were obtained through literature reviews, preliminary observations, and interviews with science teachers. The analysis also covered the Learning Outcomes (CP), Learning Objectives (TP), and indicators of problem-solving skills relevant to Grade VIII science topics, particularly work, energy, and simple machines. In the design stage, a test blueprint was developed based on the identified learning objectives and problem-solving indicators. The blueprint was used to formulate essay questions and corresponding answer keys. In addition, two supporting instruments were prepared: an expert validation questionnaire and a student response questionnaire. The development stage involved preparing and refining the assessment instrument based on the established blueprint. The resulting instrument consisted of 10 essay questions covering work, energy, and simple machines, with cognitive levels ranging from C4 to C5 based on Bloom's taxonomy (Rahayu et al., 2021). Each question was designed to assess the four stages of Polya's problem-solving framework: understanding the problem, devising a plan, carrying out the plan, and looking back. The developed instrument was subsequently submitted for expert validation and preliminary student trials. Expert validation involved four validators, while the student trials consisted of a one-to-one test involving three students and a small-group test involving ten students at SMPN 1 Ketapang, Sampang Regency.

Data collection was conducted by administering validation questionnaires and student response questionnaires. The content validity of the developed instrument was analyzed using Aiken's V coefficient based on the ratings provided by the four validators. The coefficient was calculated using the following formula in equation (1).

$$V = \frac{\Sigma s}{[n(c-1)]} \quad (1)$$

(Darmawati *et al*., 2025)

Where s is the result of r–lo, r is the number given by an assessor, lo is the lowest validation assessor number, n is the number of assessors, and c is the highest validity assessor number. Then, the appropriate assessment criteria in Table 1.

Table 1. Validity Criteria

V Value	Criteria
V ≥ 0.8	Valid
V < 0.8	Invalid

(Darmawati *et al.*, 2025)

After conducting a trial run of the test instruments, students were asked to provide feedback by completing a student response questionnaire. To calculate the percentage score from the questionnaire results, the formula listed in Equation 2 was used.

$$P = \frac{F}{N} \times 100 \quad (2)$$

(Paduwinata *et al*., 2024)

Explanation: P is the percentage, F is the total score obtained by the students, and N is the total of the indicator scores. The assessment criteria for the student response questionnaire are presented in Table 2.

Table 2. Assessment Criteria for The Student Response Questionnaire

Percentage (%)	Criteria
75 < P ≤ 100	Very good
50 < P ≤ 75	Goog
25 < P ≤ 50	Poor
0 ≤ P ≤ 25	Very Poor

(Paduwinata *et al*., 2024)

The implementation stage of the ADDIE model was not conducted in this study because the research focused on instrument development and preliminary feasibility assessment rather than classroom implementation or effectiveness testing. The one-to-one and small-group trials were conducted to obtain students' initial responses to the developed instrument. The evaluation stage was carried out formatively throughout the development process. Revisions were made based on expert feedback and student responses to improve the content, presentation, and clarity of the assessment instrument. This formative evaluation was intended to identify areas requiring improvement at each development stage (Widiastuti et al., 2024).

RESULTS AND DISCUSSION

The problem-solving instrument developed in this study focuses on the context of Toroan Waterfall and has obtained an adequate validity assessment (Valid). The instrument development process follows the stages outlined in the ADDIE model, which include analysis, design, development, implementation, and evaluation.

The first stage is analysis. Researchers identified potential science concepts found in the local wisdom of Toroan Waterfall, as well as student characteristics, learning needs, Learning Outcomes, Learning Objectives, and indicators of problem-solving skills. The researcher identified the potential science concepts contained in the local wisdom of Toroan Waterfall, as well as student characteristics, learning needs, Learning Outcomes, Learning Objectives, and problem-solving skills indicators. The results of the analysis stages are presented in Table 3.

Table 3. Results of the analysis stage

Analysis	Results
Potential Science Concepts of Toroan Waterfall	1. Effort (visitors carrying goods around the tourist area and residents fishing at Toroan Waterfall) 2. Power (Micro Hydro Water Wheel Power Plant). 3. Kinetic Energy (Toroan Waterfall appears when water falling from a height rushes downward at high speed, demonstrating the conversion of potential Energy into kinetic Energy). 4. Potential Energy (water at the top of the Toroan Waterfall cliff stores gravitational potential Energy, the magnitude of which is influenced by the mass of water, gravitational acceleration, and height). 5. Mechanical Energy (moving water, which includes the potential Energy it possesses before falling and the kinetic Energy generated from the falling water). 6. Pulley (residents still transport loads manually on steep slopes, even though the use of pulleys can make it easier for residents). 7. Wheel (Wheel in Micro Hydro Power Plant). 8. Inclined Plane (stairs and gently sloping sandy paths used by visitors and residents). 9. Lever (a tool used to clean the area around the waterfall)
Student Characteristics	1. Students often struggle to grasp abstract concepts. 2. Students still struggle with new things they have not encountered before. 3. Their problem-solving skills can also be considered low because they lack training in problem-solving. 4. Students do not yet have sufficient experience in working on problems that link science concepts to real phenomena in the surrounding environment.
Student Needs	1. They need assessment tools that can measure and train their problem-solving skills. 2. They need a learning approach that is oriented towards real contexts so that physics concepts are easier to understand. 3. HOTS (Higher-Order Thinking Skills)-based exercises are necessary. 4. There is a need for the development of authentic assessments that integrate 21st-century skills. 5. Requires teacher guidance at every stage of problem-solving.
Learning Outcomes	After studying the material on work, energy, and simple machines, students are expected to master knowledge about work and energy as well as the working principles of simple machines, including identifying various forms of energy and analyzing the use of simple machines in the surrounding environment.
Learning Objectives	1. Students can describe the variables that determine the level of success of work. 2. Students can recognize various types of energy and explain the relationship between work and energy. 3. Students can explain the mechanism of energy conversion according to their needs. 4. Students can present information related to renewable energy sources that are relevant for use in Indonesia. 5. Students can describe the advantages of using simple machines.

	6. Students can explain how several examples of simple machines commonly found in their surroundings work. 7. Students can select and adapt the type of simple machines based on the needs and problems found in the surrounding environment.
Problem-Solving Skills Indicators	1. The problem understanding stage, where the subject is expected to be able to identify known information and determine the aspects being asked in the given question. 2. The planning stage is where the subject is expected to select the appropriate formula, technique, or strategy to address the problem at hand. 3. The implementation stage, where subjects are required to execute a previously designed strategy or formula to find a solution to the problem. 4. The review stage, where subjects review and revise their answers to ensure the accuracy and correctness of their solutions.

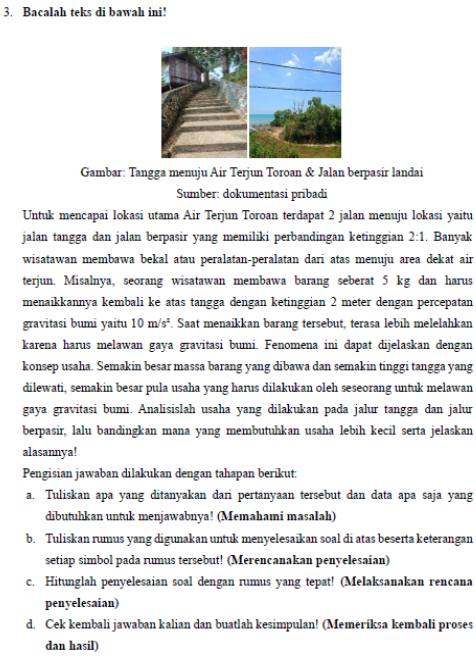
The findings suggest that Toroan Waterfall should be incorporated into science learning as a case study because it contains the concepts of work, energy, and simple machines that can be connected to other natural phenomena. Such a locality can consolidate the students' theory with the real phenomena around them (Utami & Sayekti, 2023; Dewi *et al.*, 2019). From the students' characteristics analysis, it was found that most of the students had difficulty understanding the abstract concepts, especially the work, energy and simple machines. The students have difficulty relating the concepts to real life events that happen around them and they are not used to learning that solve problems. This condition is in line with findings of Ermawan & Fauziah (2023) that concerns the understanding of physics that students have, is only on the practical activities of everyday life. Therefore, it is necessary to implement learning that has a real context and focuses on the development of higher order thinking skills (HOTS) as conceived to enrich and deepen students' understanding of the concepts (Muizz et al., 2023; Husnaini *et al*., 2025).

Based on student needs, it was found that an instrument is needed to measure and train problem-solving skills, with a context that is close to their lives. As explained by Chasanah *et al* (2022) and Murti & Sunarti (2021), the integration of local wisdom in the assessment process has been proven to encourage motivation and improve students' understanding of science material. The following analysis is Learning Outcomes (CP). Learning Outcomes are a core element of the independent curriculum, designed so that each level of education has competency standards that students must achieve. After the CP is formulated, teachers need to set Learning Objectives (TP) that serve as operational guidelines to translate the CP into concrete steps that are easy to implement in the classroom learning process (Bait *et al*., 2025). Aspects of problem-solving skills generally include several stages, namely understanding the problem, planning the solution, implementing the solution plan, and reviewing (Siswanto & Meilisari, 2024). Overall, the findings of this analysis serve as the primary reference for designing assessment instruments to measure problem-solving skills, based on the context of Toroan Waterfall in the subjects of work, energy, and simple machines. The analysis conducted serves as the foundation for designing the assessment instruments.

The second stage was the design stage, in which the findings from the analysis were translated into an assessment instrument blueprint. The blueprint was developed based on the learning objectives, science content, and four indicators of Polya's problem-solving framework. It guided the preparation of 10 essay questions, consisting of three questions on work, three on energy, and four on simple machines. The questions were designed at cognitive levels C4 (analyzing) and C5 (evaluating) according to Bloom's taxonomy (Rahayu et al., 2021). In addition to the test instrument, two supporting questionnaires were prepared. The expert validation questionnaire was designed to assess the quality of the developed questions, while the student response questionnaire was intended to gather students' perceptions after completing the instrument. The student response questionnaire covered their experiences in answering the questions, their perceptions of problem-solving-based questions compared with regular questions, and their experiences with similar questions in regular learning activities. The questionnaire was also validated before being used in the student trials.

The third stage is development. The development stage involves creating questions based on the findings from the previous stage, specifically the identification of problems (Ulviani *et al*., 2025).

Each question assesses four indicators of Polya's problem-solving skills, spanning cognitive levels C4 to C5. Questions that measure students' understanding of the concept of effort can be found in numbers 3, 8, and 9 in the developed instrument. Toroan Waterfall Tourism, which is presented in two routes to Toroan Waterfall, is illustrated in Figure 2.



3. **Bacalah teks di bawah ini!**

Gambar: Tangga menuju Air Terjun Toroan & Jalan berpasir landai
Sumber: dokumentasi pribadi

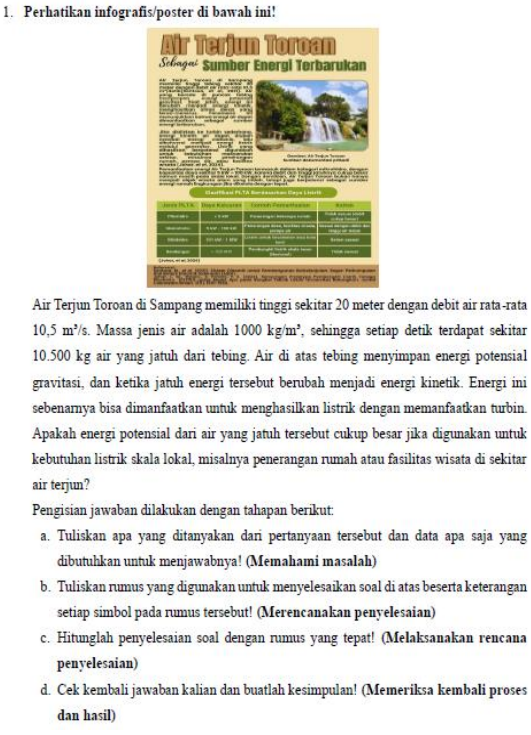
Untuk mencapai lokasi utama Air Terjun Toroan terdapat 2 jalan menuju lokasi yaitu jalan tangga dan jalan berpasir yang memiliki perbandingan ketinggian 2:1. Banyak wisatawan membawa bekal atau peralatan-peralatan dari atas menuju area dekat air terjun. Misalnya, seorang wisatawan membawa barang seberat 5 kg dan harus menaikkannya kembali ke atas tangga dengan ketinggian 2 meter dengan percepatan gravitasi bumi yaitu 10 m/s². Saat menaikkan barang tersebut, terasa lebih melelahkan karena harus melawan gaya gravitasi bumi. Fenomena ini dapat dijelaskan dengan konsep usaha. Semakin besar massa barang yang dibawa dan semakin tinggi tangga yang dilewati, semakin besar pula usaha yang harus dilakukan oleh seseorang untuk melawan gaya gravitasi bumi. Analisislah usaha yang dilakukan pada jalur tangga dan jalur berpasir, lalu bandingkan mana yang membutuhkan usaha lebih kecil serta jelaskan alasannya!

Pengisian jawaban dilakukan dengan tahapan berikut:

a. Tuliskan apa yang ditanyakan dari pertanyaan tersebut dan data apa saja yang dibutuhkan untuk menjawabnya! **(Memahami masalah)**
b. Tuliskan rumus yang digunakan untuk menyelesaikan soal di atas beserta keterangan setiap simbol pada rumus tersebut! **(Merencanakan penyelesaian)**
c. Hitunglah penyelesaian soal dengan rumus yang tepat! **(Melaksanakan rencana penyelesaian)**
d. Cek kembali jawaban kalian dan buatlah kesimpulan! **(Memeriksa kembali proses dan hasil)**

Fig.2. Excerpt from question number 3 of the problem-solving skills test instrument

Questions that measure students' understanding of the concept of energy can be found in numbers 1, 2, and 10 in the developed instrument. These questions cover four indicators of Polya's problem-solving skills at the C4-C5 cognitive level. Toroan Waterfall Tourism, which is presented in one of the questions, specifically regarding the potential energy of waterfalls, is illustrated in Figure 3.



1. **Perhatikan infografis/poster di bawah ini!**

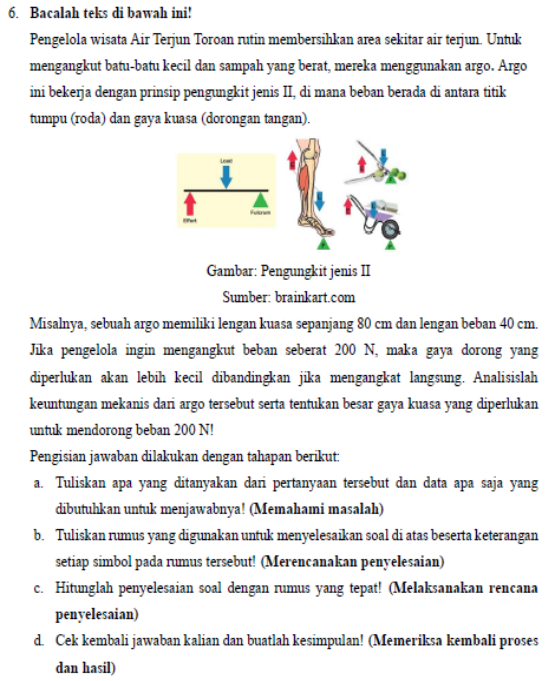
Air Terjun Toroan di Sampang memiliki tinggi sekitar 20 meter dengan debit air rata-rata 10,5 m³/s. Massa jenis air adalah 1000 kg/m³, sehingga setiap detik terdapat sekitar 10.500 kg air yang jatuh dari tebing. Air di atas tebing menyimpan energi potensial gravitasi, dan ketika jatuh energi tersebut berubah menjadi energi kinetik. Energi ini sebenarnya bisa dimanfaatkan untuk menghasilkan listrik dengan memanfaatkan turbin. Apakah energi potensial dari air yang jatuh tersebut cukup besar jika digunakan untuk kebutuhan listrik skala lokal, misalnya penerangan rumah atau fasilitas wisata di sekitar air terjun?

Pengisian jawaban dilakukan dengan tahapan berikut:

a. Tuliskan apa yang ditanyakan dari pertanyaan tersebut dan data apa saja yang dibutuhkan untuk menjawabnya! **(Memahami masalah)**
b. Tuliskan rumus yang digunakan untuk menyelesaikan soal di atas beserta keterangan setiap simbol pada rumus tersebut! **(Merencanakan penyelesaian)**
c. Hitunglah penyelesaian soal dengan rumus yang tepat! **(Melaksanakan rencana penyelesaian)**
d. Cek kembali jawaban kalian dan buatlah kesimpulan! **(Memeriksa kembali proses dan hasil)**

Fig.3. Excerpt from question number 1 of the problem-solving skills test instrument

Questions that measure students' understanding of simple aircraft concepts can be found in numbers 4, 5, 6, and 7 in the developed instrument. These questions cover four indicators of Polya's problem-solving skills at cognitive levels C4-C5. Toroan Waterfall Tourism, one of which is presented by a tourism manager who uses an argot as a Type II lever, is illustrated in Figure 4.



6. **Bacalah teks di bawah ini!**

Pengelola wisata Air Terjun Toroan rutin membersihkan area sekitar air terjun. Untuk mengangkut batu-batu kecil dan sampah yang berat, mereka menggunakan argo. Argo ini bekerja dengan prinsip pengungkit jenis II, di mana beban berada di antara titik tumpu (roda) dan gaya kuasa (dorongan tangan).

Gambar: Pengungkit jenis II

Sumber: brainkart.com

Misalnya, sebuah argo memiliki lengan kuasa sepanjang 80 cm dan lengan beban 40 cm. Jika pengelola ingin mengangkut beban seberat 200 N, maka gaya dorong yang diperlukan akan lebih kecil dibandingkan jika mengangkat langsung. Analisislah keuntungan mekanis dari argo tersebut serta tentukan besar gaya kuasa yang diperlukan untuk mendorong beban 200 N!

Pengisian jawaban dilakukan dengan tahapan berikut:

a. Tuliskan apa yang ditanyakan dari pertanyaan tersebut dan data apa saja yang dibutuhkan untuk menjawabnya! **(Memahami masalah)**
b. Tuliskan rumus yang digunakan untuk menyelesaikan soal di atas beserta keterangan setiap simbol pada rumus tersebut! **(Merencanakan penyelesaian)**
c. Hitunglah penyelesaian soal dengan rumus yang tepat! **(Melaksanakan rencana penyelesaian)**
d. Cek kembali jawaban kalian dan buatlah kesimpulan! **(Memeriksa kembali proses dan hasil)**

Fig.4. Excerpt from question number 6 of the problem-solving skills test instrument

The instrument was designed to assess students' problem-solving skills and then validated to ensure its suitability. The validation stage aims to gather input, criticism, and suggestions from validators as material for improving the product *(Lestari et al.*, 2023). The subject matter expert validation process involved two lecturers from the Science Education Study Program at Universitas Trunojoyo Madura, and two science teachers from UPTD SMPN 1 Ketapang. Validation by subject matter experts began with the involvement of two lecturers from the Science Education Study Program at Universitas Trunojoyo Madura, to ensure the suitability of scientific concepts, problem-solving skills indicators, and the accuracy of question construction in theoretical terms. After that, validation was carried out by two science teachers from UPTD SMPN 1 Ketapang to assess the comprehensibility of the language, the level of difficulty, and the relevance of the questions to the characteristics of the students at the school.

Validation was carried out on four main aspects, namely: (A) the relevance of the questions to the problem-solving skills indicators to be achieved, (B) the suitability of the questions to the answer keys based on the grid, (C) the suitability of local wisdom concepts to science, mathematics, and natural sciences material, and (D) the clarity of language use to avoid ambiguity. The assessment scale used in the validation was as follows: a score of 4 indicated "very good," a score of 3 indicated "good," a score of 2 indicated "poor," and a score of 1 indicated "very poor." The average assessment results of the validation by science subject matter experts are presented in Table 4.

Table 4. Validation Results by Science Subject Matter Experts

Item	V			
	Aspect A	Aspect B	Aspect C	Aspect D
1	0,88	1,00	1,00	0,88
2	1,00	0,94	1,00	0,88
3	0,94	0,88	1,00	1,00
4	1,00	0,94	1,00	0,88

Item	V			
	Aspect A	Aspect B	Aspect C	Aspect D
5	1,00	0,94	1,00	0,94
6	0,94	1,00	1,00	1,00
7	0,94	0,88	1,00	1,00
8	0,94	1,00	1,00	1,00
9	1,00	1,00	0,94	1,00
10	0,94	1,00	1,00	1,00
Total	0,96	0,96	0,99	0,96
Average	0.96			
Category	Highly valid			

The data in Table 4 provide a summary of the average results from the validation analysis conducted by subject matter experts. The validation assessment results for aspect A show an average score of 0.96, indicating that the questions developed are aligned with the expected indicators of problem-solving skills. Aspect B also received an average score of 0.96, indicating the relevance of the questions to the answer key. Aspect C received an average score of 0.99, indicating the relevance of local wisdom concepts and science material used in the questions. Aspect D received an average score of 0.96, verifying that the language of the questions was clearly stated without the possibility of cross interpretation. Furthermore, instrument validation overall average score was 0.96, indicating it was highly valid to be used in the context of Toroan Waterfall for work, energy, and simple machines. Validators based on questionnaires also claimed that this instrument was valid and applicable in the field with a few adjustments recommended. As stated by Maulana (2022) explains, validity testing, is a form of determination that examines whether the tool or instrument in question can indeed measure or represent the trait to be observed. A comparable result was reported in Ulviani *et al* (2025) when they validated the physics problem-solving instrument they constructed with Aiken's V and got an average of 0.836, which is in the category of very high validity.

Once the student response questionnaire has been completed, the next stage was to validate it with the science lecturers of Universitas Trunojoyo Madura prior to the one-on-one and the small group testing. The results of the response questionnaire validation showed a score of 0.81, placing it in the category of valid instruments. This goes hand in hand with Darmawati *et al* (2025) where it was found that in order for the instruments in the questionnaire to measure students' reaction to the developed product with accuracy and dependability, the questionnaire instruments have to undergo validity testing first. Once the validation was complete, the problem-solving questions and the students' response to the questionnaire were trialed to students using the one to one and small group methods. The response students provided from the questionnaire is captured in Table 5.

Table 5. Results of Student Response Questionnaires

x	One-to-One Test	Small Group Test
Average	72%	73%
Category	Good	Good

Table 5 presents the results of the student response questionnaire assessment, where the average result of the one-to-one test was 72%, which is categorized as good, and the small group test was 73%, which is also categorized as good. These findings indicate that the instrument for assessing problem-solving skills in the context of Toroan Waterfall, energy, work, and simple machines was well received by students, suggesting that it was both practical and enjoyable for them. These findings align with the results of research conducted by Ulviani *et al* (2025), which reported that the average percentage of positive responses from students to the test reached 82% in the "outstanding" category. This means that the problem-solving skills instrument developed received a positive response from the majority of students and can be used by teachers to assess problem-solving skills in the learning process.

The fourth stage, product implementation, was not carried out in this study; therefore, the process proceeded directly to the next stage, evaluation. The evaluation phase, which is the final phase of the ADDIE model, is carried out to determine the usefulness of the instruments using actual data

gathered in the implementation phase (Putra *et al*., 2025). There has been an ongoing formative evaluation even during the analysis, design and development phases which is in sustain of using the product to narrow it down for wider use (Ulviani *et al*., 2025). There was a formative evaluation in every of the phases in order to allow the product to be fine-tuned rapidly based on the feedback *(*Widiastuti *et al*., 2024). The validation phase of the instruments also had the experts provide recommendations and comments that were then used to make further modifications and improvements to the instruments to enhance their usability.

CONCLUSION AND SUGGESTION

This study resulted in a problem-solving skills assessment instrument consisting of 10 essay questions covering work, energy, and simple machines in the context of Toroan Waterfall. The questions were developed based on the four stages of Polya's problem-solving framework: understanding the problem, devising a plan, carrying out the plan, and looking back. The expert validation results yielded an average Aiken's V coefficient of 0.96, indicating that the instrument met the established content validity criteria. Furthermore, the student response questionnaire results showed percentages of 72% in the one-to-one trial and 73% in the small-group trial, both of which were categorized as good. These findings indicate that the developed instrument has met the content validity criteria and received positive responses from students during the preliminary trials. However, this study was limited to instrument development, expert validation, and preliminary student trials. The classroom implementation stage was not conducted; therefore, the instrument's applicability in actual learning activities and its effectiveness in assessing students' problem-solving skills have not yet been examined. Future research is recommended to conduct classroom implementation involving a larger and more diverse group of students. Further studies should also examine the instrument's applicability in science learning and its ability to measure students' problem-solving skills in broader learning contexts.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

AUTHOR CONTRIBUTIONS STATEMENT

Conceptualization, L.Q. and H.M.; methodology, L.Q., H.M., and W.P.H.; investigation, L.Q.; data curation, L.Q.; formal analysis, L.Q. and H.M.; writing—original draft preparation, L.Q.; writing—review and editing, H.M. and W.P.H.; supervision, H.M. and W.P.H. All authors have read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

This research was supported by the Institute for Research and Community Service (LPPM) of Universitas Trunojoyo Madura with research contract number No. 152/UN46.4.1/PT.01.03/RISMAN/2025.

DECLARATION OF GENERATIVE AI SOURCES

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) solely for language editing, grammar checking, and improving the clarity of the English text. All content was reviewed and verified by the authors, who take full responsibility for the final manuscript.

REFERENCES

Amiruddin, Rochman, C., & Nana. (2024). Mengukur efektivitas model pembelajaran berbasis masalah (PBL) dalam pembelajaran IPA. *Jurnal Pendidikan MIPA, 14*(3), 723–731. https://doi.org/10.37630/jpm.v14i3.1790

Bait, E. H., Mulyasari, E., Hendriawan, D., Arwasih, & Ulwan, M. N. (2025). Kurikulum Merdeka dan dinamika tujuan pendidikan: Integrasi capaian pembelajaran (CP), tujuan pembelajaran (TP), dan alur tujuan pembelajaran (ATP). *Kalam Cendekia: Jurnal Ilmiah Kependidikan, 13*(1), 615–624. https://doi.org/10.20961/jkc.v13i1.97505

Chasanah, N., Widodo, W., & Suprapto, N. (2022). Pengembangan instrumen asesmen literasi sains untuk mendeskripsikan profil peserta didik. *PENDIPA Journal of Science Education, 6*(2), 474–483. https://doi.org/10.33369/pendipa.6.2.474-483

Darmawati, V. D., Wulandari, A. Y. R., Putra, D. B. R. A., & Sulastri, E. (2025). Validitas instrumen soal literasi sains pada konteks batik Madura materi unsur, senyawa, dan campuran. *Jurnal Pendidikan MIPA, 15*(1), 173–183. https://doi.org/10.37630/jpm.v15i1.2284

Dewi, C. A., Khery, Y., & Erna, M. (2019). An ethnoscience study in chemistry learning to develop scientific literacy. *Jurnal Pendidikan IPA Indonesia, 8*(2), 279–287. https://doi.org/10.15294/jpii.v8i2.19261

Ermawan, M. Z. F., & Fauziah, A. N. M. (2023). Analisis kemampuan pemecahan masalah IPA pada siswa SMP dalam menyelesaikan soal IPA. *Lentera: Multidisciplinary Studies, 1*(2), 75–82. https://lentera.publikasiku.id/index.php/ln/article/view/11

Herawati, S. S., Kurniawan, D., & Rahmanita, U. (2024). Pengembangan video animasi berbasis Animaker menggunakan model ADDIE pada topik karakteristik materi dan perubahannya. *Jurnal Pendidikan Ilmu Pengetahuan Alam (JP-IPA), 5*(2), 141–151. https://doi.org/10.56842/jp-ipa.v5i2.383

Husnaini, D., Darmaji, D., & Kurniawan, D. A. (2025). Meningkatkan kemampuan berpikir kreatif siswa melalui model pembelajaran discovery learning pada materi usaha, energi dan pesawat sederhana di kelas VIII SMP Negeri 22 Kota Jambi. *JIIP: Jurnal Ilmiah Ilmu Pendidikan, 8*(4), 4129–4133. https://doi.org/10.54371/jiip.v8i4.7637

Khaerunnisa, E., Muhyidin, A., & Leksono, S. M. (2025). Instrumen kemampuan pemecahan masalah bangun ruang sisi datar pada konteks rumah adat Baduy. *Jurnal Penelitian dan Pembelajaran Matematika, 18*(2), 107–123. https://doi.org/10.30870/jppm.v18i2.33947

Krismawati, A., Maula, L. H., & Prayogo, M. S. (2023). Pengembangan media pembelajaran IPA berbasis papan pintar pada materi peran fungi dalam kehidupan sehari-hari di kelas IV SD. *Jurnal Tadris IPA Indonesia, 3*(2), 219–227. https://doi.org/10.21154/jtii.v3i2.1745

Lestari, R., Jasiah, Rizal, S. U., & Syar, N. I. (2023). Pengembangan media berbasis video pada pembelajaran IPAS materi permasalahan lingkungan di kelas V SD. *HOLISTIKA: Jurnal Ilmiah PGSD, 7*(1), 34–43. https://doi.org/10.24853/holistika.7.1.34-43

Maulana, A. (2022). Analisis validitas, reliabilitas, dan kelayakan instrumen penilaian rasa percaya diri siswa. *Jurnal Kualitas Pendidikan, 3*(3), 133–139. https://doi.org/10.51651/jkp.v3i3.331

Muizz, A., Suryanti, & Prahani, B. K. (2023). Literature review: Penggunaan modul IPA berbasis etnosains untuk meningkatkan literasi sains pada siswa SD. *Jurnal Elementaria Edukasia, 6*(4), 1905–1914. https://doi.org/10.31949/jee.v6i4.7574

Murti, W. W., & Sunarti, T. (2021). Pengembangan instrumen tes literasi sains berbasis kearifan lokal di Trenggalek. *ORBITA: Jurnal Kajian, Inovasi dan Aplikasi Pendidikan Fisika, 7*(1), 33–43. https://doi.org/10.31764/orbita.v7i1.4386

Nirmalasari, M. A. Y., Mago, O. Y. T., & Manuk, I. L. (2022). Validitas instrumen soal literasi numerasi kimia hidrokarbon dalam integrasinya dengan isu sosiosaintifik lokal Sikka. *Jurnal Pendidikan MIPA, 12*(4), 1004–1011. https://doi.org/10.37630/jpm.v12i4.725

Nurfadillah, T., Elvia, R., & Elvinawati. (2023). Pengembangan instrumen tes kimia berbasis literasi sains untuk mengukur literasi sains siswa. *ALOTROP: Jurnal Pendidikan dan Ilmu Kimia, 7*(1), 44–56. https://doi.org/10.33369/alo.v7i1.28253

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

Paduwinata, V., Wulandari, A. Y. R., Tamam, B., Wahyuni, E. A., & Hartiningsih, T. (2024). Development of science literacy-based e-booklet accompanied by mind mapping on human digestive system material. *Journal of Education and Learning Research, 2*(1), 19–27. https://doi.org/10.62208/jelr.2.1.p.19-27

Polya, G. (1957). *How to solve it: A new aspect of mathematical method* (2nd ed.). Princeton University Press.

Putra, D. D., Nindiasari, H., & Fathurrohman, M. (2025). Pengembangan instrumen tes kemampuan pemecahan masalah matematis siswa sekolah menengah kejuruan. *Edukasiana: Jurnal Inovasi Pendidikan, 4*(3), 967–983. https://doi.org/10.56916/ejip.v4i3.1610

Rahayu, L. S., Susanta, A., & Yensy, N. A. (2021). Analisis soal pada buku sekolah elektronik (BSE) pokok bahasan teorema Pythagoras berdasarkan taksonomi Bloom. *Jurnal Penelitian Pembelajaran Matematika Sekolah (JP2MS), 5*(1), 50–57. https://doi.org/10.33369/jp2ms.5.1.50-57

Sari, Y. K., & Tanjung, S. (2022). *Meta analisis terhadap pengaruh pembelajaran inkuiri dalam pemecahan masalah matematis siswa SLTP*. Edu Publisher.

Siswanto, E., & Meilisari. (2024). Penerapan metode pembuktian dalam logika matematika. *JRPMS: Jurnal Riset Pembelajaran Matematika Sekolah, 8*(1), 112–118. https://doi.org/10.21009/jrpms.081.06

Tanjung, Y. I., Wulandari, T., & Festiyed. (2024). Pengembangan instrumen tes kemampuan pemecahan masalah dalam konteks budaya pada materi mekanika. *Naturalistic: Jurnal Kajian dan Penelitian Pendidikan dan Pembelajaran, 8*(2), 339–345. https://doi.org/10.35568/naturalistic.v8i2.3514

Ulviani, N., Maria, H. T., & Hidayatullah, M. M. S. (2025). Pengembangan instrumen tes kemampuan pemecahan masalah peserta didik pada hukum Newton di SMA Negeri 1 Teluk Pakedai. *Jurnal Education and Development, 13*(1), 100–106. https://doi.org/10.37081/ed.v13i1.6389

Utami, A. T., & Sayekti, I. C. (2023). Kajian etnosains pembuatan pisau sebagai kearifan lokal Kabupaten Klaten pada materi pembelajaran IPA sekolah dasar. *Attadib: Journal of Elementary Education, 7*(2). https://doi.org/10.32832/at-tadib.v7i2.19505

Utami, H. S., & Puspitasari, N. (2022). Kemampuan pemecahan masalah siswa SMP dalam menyelesaikan soal cerita pada materi persamaan kuadrat. *Jurnal Inovasi Pembelajaran Matematika: PowerMathEdu, 1*(1), 57–68. https://doi.org/10.31980/powermathedu.v1i1.1916

Widiastuti, W., Linora, L., & Hidayat, H. (2024). Menganalisis pemilihan model evaluasi program pendidikan. *Jurnal Penelitian Ilmu Pendidikan Indonesia, 3*(3), 383–394. https://doi.org/10.31004/jpion.v3i3.299