

Research Article

MAPACCING: Development of a STEM–Design Thinking Physics Learning Module Based on Local Materials to Improve the Problem-Solving Skills of Coastal Senior High School Students in Majene

Sulfianty^{1*}, Andi Kamal Ahmad¹, Khairil Asnan Haedar², Jupri, Mina¹

¹STKIP Darud Dawah Wal Irsyad Pinrang, Pinrang, Indonesia

²Universitas Sibatokong Mambo, Bone, Indonesia

*Vivisulfianty27@gmail.com

ABSTRACT

Coastal water quality problems are very close to the everyday lives of students in Majene Regency, yet they have not been optimally used as a context for physics learning to develop students' problem-solving skills. This study aimed to develop the MAPACCING Learning Module, based on STEM–Design Thinking and using local materials, and to determine its validity, practicality, and initial effectiveness in improving students' problem-solving skills. The study used the Research and Development (R&D) method with the 4D model, consisting of the define, design, develop, and disseminate stages. The initial field effectiveness test used a nonequivalent pretest–posttest control group quasi-experimental design involving 22 students in the experimental group and 20 students in the control group at SMA Negeri 3 Majene. The results showed that the module was valid (moderate category), with Aiken's V coefficients ranging from 0.72 to 0.79. The mean student response score was 116.6 (good), while the practitioner response score was 120.67 (very good). The problem-solving skills instrument consisted of 12 valid items with a reliability coefficient of 0.72. The mean N-gain was 0.42 (moderate) in the experimental group and 0.29 (low) in the control group. An independent-samples t-test showed a significant difference in the improvement of problem-solving skills between the experimental and control groups ($t = 5.09$, $p < 0.001$), with a Cohen's d of 1.57. These findings provide initial evidence of the effectiveness of the MAPACCING module in supporting students' problem-solving skills through STEM–Design Thinking physics learning contextualized in coastal water quality and local materials.

Keywords: Design Thinking; Local Materials; Problem-Solving Skills; Physics Education; STEM.

ARTICLE HISTORY

Received: 03.09.2026

Accepted: 14.10.2026

Published: 21.10.2026

ARTICLE LICENCE

Copyright ©2026 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike 4.0 International (CC BY-SA 4.0)

1. Introduction

Majene Regency is a region in West Sulawesi Province with a coastline of approximately 125 km (Sunarwan & Firmansyah, 2016). This coastal setting makes seawater intrusion and groundwater quality relevant environmental concerns for communities in the region. Seawater intrusion can cause groundwater in residents' wells to have high salinity (a salty taste), turbidity, and odor, so the availability of clean water suitable for domestic needs is very limited (Zahra et al., 2022). Declining groundwater quality and health risks from contamination of coastal aquifers have also been reported in other regions (Yassin et al., 2024). This condition is not merely a global environmental issue but an everyday contextual problem that is very close to the lives of students in Majene Regency.

In the context of education, this water quality problem has great potential to serve as a meaningful learning resource. This is in line with the spirit of the Merdeka Curriculum, which emphasizes contextual learning, the strengthening of 21st-century competencies, and the connection between academic concepts and real issues in students' surroundings (Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, 2022). Kusnandar (2025) developed coastal environment-based science learning for Madrasah Tsanawiyah students, which shows the potential of this context, but evidence at the level of physics learning in senior high schools remains limited.

The opportunity for educational transformation based on local issues is reinforced by the study of Sulfianty et al. (2024) at SMAN 1 Sendana. Their findings reported that STEM-based learning materials linked to the surrounding environment were considered valid, practical, and effective in improving student learning outcomes. This finding suggests the feasibility of contextual STEM-based learning in at least one school in Majene Regency, although further testing in other school contexts is still needed. However, that study still focused on strengthening the understanding of basic concepts and had not yet reached the stage of designing a water filtration technology prototype.

This gap is the focus of the present study. On the one hand, Majene has the potential of local materials such as coconut husk, coconut shell, and beach sand, which in various earlier studies have shown potential for use as simple water filtration media and adsorbents (Awang et al., 2025; Che Osmi et al., 2024). Adsorbent materials have also been studied for removing oil contaminants from water (De Sá et al., 2019). On the other hand, based on interviews with teachers during the define stage, teachers do not yet have systematic guidance for integrating the potential of these materials into learning as contextual engineering projects. This condition may limit students' opportunities to practice problem-solving skills through authentic engineering experiences (Syukri et al., 2018), even though learning based on the surrounding environment can develop students' science process skills and problem-solving abilities (Wahyuni et al., 2017).

To meet this need, the MAPACCING Learning Module was developed. The name, taken from Mandar local wisdom and meaning "Clean," combines the STEM approach with the *Design Thinking* framework (Razzouk & Shute, 2012). The module is designed to be semi-independent and is equipped with usage instructions, learning objectives, contextual material, STEM project activities, and formative evaluation. The feasibility of the module was tested through expert validation of its content, media, and language using Aiken's V formula, and its practicality was tested through practitioner and student responses regarding the content, presentation, language, and graphics of the module. Developing modules with expert validation and user trials is a common practice for obtaining suitable teaching materials (Hasanah et al., 2025; Pagiling et al., 2024).

The learning activities in the module are carried out through the design of a simple water filter prototype that uses local materials as a contextual learning resource. These activities focus on engineering thinking processes and simple data analysis through the measurement of water quality parameters, such as total dissolved solids (TDS), pH, and turbidity, as a means of training students' problem-solving skills. In addition, the learning integrates physics concepts related to the physical properties of matter and fluids, such as fluid flow through filtering media and changes in the physical properties of water before and after filtration.

Previous research has demonstrated the benefits of each of these components separately. STEM-based learning materials linked to the surrounding environment have been developed with a focus on student learning outcomes (Sulfianty et al., 2024); the integration of *Design Thinking* into STEM learning has been associated with the problem-solving skills of senior high school students (Nguyễn et al., 2025); the engineering design

process in science learning has been examined as an effort to improve problem-solving skills (Syukri et al., 2018; Winarno et al., 2020); and local materials such as coconut shell have been studied as water filters or adsorbents in engineering and materials science (Awang et al., 2025; Che Osmi et al., 2024). However, based on the literature reviewed in this article, no study appears to have combined all four elements, namely STEM, *Design Thinking*, local materials, and the measurement of problem-solving skills, in a single physics learning module that was developed systematically, validated by experts, tested for practicality, and tested for initial effectiveness among students in a coastal area. The novelty of this study lies in integrating these four elements into the MAPACCING Module within the context of coastal water quality in Majene, using local materials as resources for engineering-design activities, as well as in providing initial empirical evidence on product quality and effectiveness, expressed as differences in problem-solving skills associated with the use of the module rather than as evidence of a causal impact.

The research questions in this study are: (1) To what extent is the developed MAPACCING module valid and practical for implementation with senior high school students in Majene Regency? (2) What initial evidence is there of the effectiveness of the MAPACCING module in supporting senior high school students' problem-solving skills through STEM–Design Thinking learning in the context of coastal water quality?

2. Method

2.1 Type and Design of the Research

This study is a Research and Development (*R&D*) study that aims to produce a MAPACCING learning module that is valid, practical, and has initial effectiveness. The development procedure used the 4D model, namely the define, design, develop, and disseminate stages. The dissemination stage was carried out on a limited basis in accordance with the scope of the study.

The 4D model was used as the module development procedure, from identifying needs to producing a module that had undergone validation and trials. Subsequently, to test the initial effectiveness of the module, a nonequivalent pretest–posttest control-group quasi-experimental design using two intact classes was employed. The partner school was selected using purposive sampling based on considerations in line with the research objectives. This design involved an experimental group and a control group drawn from classes that had already been formed. The two intact classes were randomly assigned to the experimental and control conditions. Thus, randomization was conducted at the class level rather than at the individual student level. With only two classes, class-level randomization does not guarantee that the groups were equivalent at the outset. Students remained in their respective classes and were not moved or randomly assigned individually to new groups. Both groups were given a pretest before the instruction and a posttest after the instruction. The experimental group received instruction using the MAPACCING module, while the control group received instruction without using the MAPACCING module.

2.2 Research Location and Subjects

The study was conducted at SMA Negeri 3 Majene (State Senior High School 3 Majene), Majene Regency, West Sulawesi, as the partner school implementing the study. This school was selected purposively by considering access for conducting the research, the readiness of learning facilities, the support of the partner school, the closeness of students' experience to the coastal environment of Majene Regency, and the relevance of the coastal environmental context to the water quality problem

addressed in the module. Majene Regency has a coastline of approximately 125 km (Sunarwan & Firmansyah, 2016) and its territory stretches along the coast, so students at the partner school live in a coastal environment. Preliminary observations of water quality problems were carried out in several coastal settlements in Majene Regency to obtain the problem context used in developing the module. Preliminary observations were also conducted with students at the partner school to identify their initial experiences related to water quality issues in the coastal environment.

The research subjects consisted of three groups. First, three expert validators were involved to assess the feasibility of the MAPACCING module and the research instruments. The validators were selected purposively based on their educational background, academic rank, and expertise relevant to the product being developed, covering STEM content, learning media design, and language.

Validator 1 holds a master's degree in physics education and is a lecturer with the academic rank of Lektor (Assistant Professor equivalent), with expertise in STEM-based physics learning and seven years of teaching and research experience. Validator 2 holds a master's degree in physics education and is a lecturer with the academic rank of Lektor, with expertise in learning media and graphic design and seven years of professional experience. Validator 3 holds a master's degree in English Language Education and is a lecturer with the academic rank of Lektor, with 12 years of teaching experience. Validator 3 was appointed as the language expert based on his or her academic background and experience in the fields of language and education. The language assessment covered clarity, readability, grammar, and punctuation. Each validator assessed the module and the research instruments according to his or her own area of expertise.

Second, three physics/science teachers (hereinafter referred to as practitioners) were involved to assess the practicality of the MAPACCING module. The practitioners were selected purposively based on their experience teaching physics at the senior high school level and their suitability to the characteristics of the product being developed. The three practitioners reviewed the module, examined its content and learning activities, and observed the implementation of learning using the MAPACCING module during the research. Based on the results of this review and observation, the practitioners assessed the practicality of the module for implementation in learning.

Third, the field test subjects consisted of two Grade XI Physics specialization classes at SMA Negeri 3 Majene, with a total of 42 students: 22 students in the experimental class, namely XI Physics 2, who received instructions using the MAPACCING module. Meanwhile, the control class, namely XI Physics 1, consisted of 20 students who received conventional instruction (lecture, discussion, and practice exercises) without using the MAPACCING module.

This study used two Grade XI Physics specialization classes that had been formed by the school. The assignment of the experimental class and the control class were done randomly between these two classes. The result of the randomization designated class XI Physics 2 as the experimental group, consisting of 22 students, and class XI Physics 1 as the control group, consisting of 20 students. Randomization was carried out only to determine the role of each class in the study, not to assign students individually to groups. Therefore, this study used a nonequivalent pretest–posttest control-group design with class-level random assignment of two intact classes.

Research permission was obtained through a research permit application from the Institute for Research and Community Service (LPPM) of STKIP Darud Da'wah wal Irsyad Pinrang (No. 013/LPPM/STKIP-DDI/VII/2026) and a research permit letter issued by the Investment and One-Stop Integrated Services Office (DPM-PTSP) of Majene

Regency (No. 500.16.7.2/430/IP/VII/2026). The study was conducted at SMA Negeri 3 Majene after obtaining permission from the school. Before data collection, students were informed about the purpose and procedures of the study. The research data were kept confidential, and students' identities were not disclosed in reporting the results of the study.

2.3 Development Procedure

The development procedure of the MAPACCING module used the 4D model, consisting of the define, design, develop, and disseminate stages.

a) Define Stage

This stage was carried out to identify the needs for module development through a literature review, field observation, teacher interviews, analysis of student characteristics, analysis of students' initial experiences related to coastal environmental issues, and analysis of the concept of water filtration from a STEM–*Design Thinking* perspective.

b) Design Stage

This stage included preparing the initial prototype of the module and the research instruments, including developing the problem-solving skills test, designing the module structure, developing STEM–*Design Thinking*-based learning activities, and integrating local materials as a context for engineering learning.

c) Develop Stage

This stage included expert validation, empirical testing of the instruments, a field test, and product revision. The module draft and all instruments were validated by three expert validators. After being declared valid, the problem-solving skills test instrument was then tested empirically on students outside the research classes, namely classes XII Physics 1 and XII Physics 2, to determine the validity and reliability of the test items.

The field test was carried out in two classes, namely class XI Physics 2 as the experimental group and class XI Physics 1 as the control group, using a *Nonequivalent Control Group Pretest–Posttest* design. The field test was carried out over six meetings. At the start of the study, both groups were given a *pretest* to determine their initial problem-solving skills. Next, the experimental group took part in learning activities using the MAPACCING module based on STEM–*Design Thinking*, while the control group took part in learning activities without using the MAPACCING module. After all learning activities were completed, both groups were given a *post test* to determine the change in their problem-solving skills after the learning process.

Learning conditions in both groups were kept equivalent. Both groups were taught by the research team, and each meeting had the same duration in both groups. In the experimental group, the MAPACCING module was used over six meetings (Learning Activities 1–6, KB1–KB6), with learning activities following the stages of Empathize, Define, Ideate, Prototype, Test, and Communication. The control group studied the same material, namely the concepts of static and dynamic fluids. Instruction in the control group used conventional teaching methods, including lecture, discussion, and practice exercises. The main activities in the control class included explaining the material, presenting example problems, practice exercises, and summaries. Learning activities were carried out with the same number of meetings in both groups to ensure equivalent learning time allocation. The same *pretest* and *posttest* were given to both groups, with identical test duration, scoring rubrics, and scoring procedures. The *post test* was administered after all learning meetings had been completed.

d) Disseminate Stage

This stage was carried out on a limited basis by distributing the MAPACING module product at the partner school because of the development.

2.4 Research Instruments

The research instruments used included a module validation sheet, an instrument validation sheet, a student response questionnaire, a practitioner response questionnaire, and a problem-solving skills test. The instruments were developed based on the research objectives, the characteristics of the MAPACING module, and indicators relevant to the aspects being measured.

a) Module Validation Sheet

The module validation sheet was used to obtain expert assessments of the feasibility of the MAPACING module. This instrument consisted of 66 items covering the aspects of content feasibility, presentation, language, and graphics. The assessment used a Likert scale (scale of 1–4), with scores indicating the degree of suitability or feasibility of each aspect. The validation sheet was prepared based on criteria for developing teaching materials and the characteristics of STEM–*Design Thinking*-based modules.

b) Instrument Validation Sheet

The instrument validation sheet was used to assess the suitability of the research instruments before they were used for data collection. The aspects assessed included content feasibility, presentation, language, and graphics/media, with a total of 40 items. The assessment used a scale of 1–4. This instrument was developed based on the indicators of each instrument and was used as the basis for making improvements before the instruments were used in the research stage.

c) Student Response Questionnaire

The student response questionnaire was initially developed with 40 statements covering the dimensions of content feasibility, presentation, language, graphics, and instructions. After expert validation, the instructions dimension was removed, resulting in a final questionnaire consisting of 37 statements. The final questionnaire covered four dimensions: content feasibility (13 items), presentation (5 items), language (5 items), and graphics (14 items), for a total of 37 items. The items assessing students' responses to the problem-solving aspects of the MAPACING module were included in the content feasibility dimension. The questionnaire used a four-point Likert scale (1–4) and was developed based on the characteristics of the module and the practicality aspects to be examined from the students' point of view.

d) Practitioner Response Questionnaire

The practitioner response questionnaire was used to assess the practicality of the MAPACING module based on the practitioners' review of the module and their observations during classroom learning using the module. After observing all learning meetings, the practitioners assessed the module on the aspects of content feasibility, presentation, language, and graphics. The questionnaire consisted of 34 statements covering four aspects, namely content feasibility (12 items), presentation (7 items), language (6 items), and graphics (9 items). The assessment used a four-point Likert scale (1–4). The questionnaire was developed based on indicators for assessing product practicality and the characteristics of the module for use in physics learning.

e) Problem-Solving Skills Test

The problem-solving skills test was used to determine changes in students' problem-solving skills before and after instruction. The test was given in the form of 12 essay items at the *pretest* and *posttest* stages. The items were developed based on four indicators of problem-solving skills: understanding the problem, planning the solution, carrying out the solution, and evaluating or checking the solution.

The test was scored using an analytic rubric developed based on these four indicators of problem-solving skills. The twelve essay items represented the four indicators, with three items representing each indicator. Each item was scored on a scale of 0–3 using the rubric descriptors of the indicator it represented, according to the level of achievement shown in the student's answer. The scores obtained from the twelve items were summed to obtain the total test score, with a minimum score of 0 and a maximum score of 36. The same scoring procedure was applied to the *pretest* and *posttest* to measure changes in students' problem-solving skills after the learning process.

2.5 Validity Analysis

The expert validation analysis was carried out to determine the level of item relevance based on the assessments given by three validators, using Aiken's content validity coefficient (Aiken's V). Each validator rated each item using a rating scale of 1–4, from the lowest to the highest level of relevance. The analysis used Aiken's content validity coefficient with the following formula (Aiken, 1985; Azwar, 2012; Kania et al., 2024).

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

Where $s = r - I_0$, (r) denotes the rating given by the validator, (I_0) denotes the lowest rating on the scale, (n) denotes the number of validators, and (c) denotes the number of rating categories. Aiken's V ranges from 0 to 1, with higher values indicating greater agreement among validators regarding item relevance. The three validators rated each item on a four-point scale ($c = 4$). Aiken (1985) provides minimum V values for statistical significance that depend on the number of raters and the number of rating categories; with only three raters, these values are very high, so significance-based thresholds were not used. Instead, a value of $V \geq 0.40$ was used as the minimum practical criterion for acceptable content validity, and V values were interpreted descriptively using the categories proposed by Retnawati (2016): low ($V < 0.40$), moderate ($0.40 \leq V \leq 0.80$), and high ($V > 0.80$). Thus, V here is interpreted as an index of expert agreement and not as a statistical significance test, and the results are reported by category, not merely as pass or fail. Aiken's V values were then used to describe the level of content validity of the instruments based on the validators' assessments.

2.6 Problem-Solving Skills Test

The items of the problem-solving skills test that had been validated by experts were tested empirically on students outside the research classes. The empirical test was conducted with 45 Grade XII Physics specialization students who had characteristics relatively like those of the research subjects and who had studied the material measured by the test. The empirical test was conducted to determine the validity and reliability of the test items before they were used at the *pretest* and *posttest* stages.

Item validity analysis of the essay test used the *Pearson Product Moment* correlation between each item score and the total test score. The reliability of the

instrument was analyzed using the *Cronbach's Alpha* coefficient to determine the internal consistency of the instrument (Arikunto, 2010).

The test was scored using analytic rubric prepared based on the achievement of each indicator in students' answers. Each item was scored from 0 to 3 according to the level of achievement of its indicator as shown in the answer. The test consisted of 12 items, so the maximum total score was 36. The final score was obtained by summing the scores of all test items. The same rubric was used to score the *pretest* and *posttest* so that the measurement results could be compared consistently.

2.7 Practicality Analysis

The practicality of the module was analyzed based on student and practitioner responses. The response data were analyzed descriptively by summing each respondent's scores across all statement items to obtain a total score, which was then grouped into practicality categories. The practicality assessment was based on students' responses after using the MAPACING module and on the practitioners', assessment based on their review of the module and their observation of learning over six meetings. Practicality categories were determined using five categories, namely very poor, poor, fair, good, and very good. The assessment of responses to the appearance, content, and usefulness of the module refers to research on STEM-based modules (Utami et al., 2023).

Category intervals were determined based on the ideal total score of the instrument. The minimum and maximum scores were calculated by multiplying the number of statement items by the lowest and highest scores on the 4-point rating scale. The interval width was then calculated using the following formula:

$$I = \frac{S_{max} - S_{min}}{5}$$

where (I) is the interval width, S_{max} is the maximum ideal score, S_{min} is the minimum ideal score, and 5 denotes the number of practicality categories.

For student responses, the instrument consisted of 37 statement items with a 4-point rating scale. Therefore, the minimum ideal score is:

$$S_{min} = 37 \times 1 = 37$$

and the maximum ideal score is:

$$S_{max} = 37 \times 4 = 148$$

Thus, the interval width is:

$$I = \frac{148 - 37}{5} = 22.2$$

Based on this calculation, the categories of student responses were determined using the total score intervals presented in Table 1. The interval width was calculated as $p = r/k$ (Sudjana, 1992; Khaeruddin & Ali, 2012), where r is the range and k is the number of classes. Because the total score is an integer, $p = 111/5 = 22.2$ was rounded up to 23. Following Khaeruddin and Ali (2012), the class boundaries were then checked using $p \cdot k = (r + 1) + X$, where X is the determining value: $23 \times 5 = (111 + 1) + X$, so $X = 3$. X was divided into 1 and 2, so that the lowest class begins at $37 - 1 = 36$ and the highest class ends at $148 + 2 = 150$.

Table 1. Student Response Categories

Class Interval	Category
128 – 150	Very good
105 – 127	Good
82 – 104	Fair
59 – 81	Poor
36 – 58	Very poor
Total	

For practitioner responses, the instrument consisted of 34 statement items with a 4-point rating scale. Therefore, the minimum ideal score is:

$$S_{min} = 34 \times 1 = 34$$

and the maximum ideal score is:

$$S_{max} = 34 \times 4 = 136$$

Thus, the interval width is:

$$I = \frac{136 - 34}{5} = 20.4$$

Accordingly, the categories of practitioner responses were determined based on the total score intervals presented in Table 2. In the same way, $p = 102/5 = 20.4$ was rounded up to 21, and $21 \times 5 = (102 + 1) + X$ yields $X = 2$. X was divided into 1 and 1, so the classes run from $34 - 1 = 33$ to $136 + 1 = 137$ (Khaeruddin & Ali, 2012).

Table 2. Practitioner Response Categories

Class Interval	Category
117 – 137	Very good
96 – 116	Good
75 – 95	Fair
54 – 74	Poor
33 – 53	Very poor
Total	

2.8 Initial Effectiveness Analysis

The initial effectiveness of the module was analyzed based on changes in students' problem-solving skills before and after instruction through *pretest* and *posttest* scores. Each student's improvement was calculated using the *N-gain*, which describes the magnitude of improvement relative to the students' initial ability. Next, the difference in improvement in problem-solving skills between the experimental group and the control group was analyzed using an independent-samples t-test on the *N-gain* scores. The use of this test was based on the purpose of the analysis, namely to determine whether there is a difference in the mean improvement in problem-solving skills between two independent groups.

The *N-gain* value was calculated using the following formula (Hake, 1998):

$$N_{gain} = \frac{\text{Posttest score} - \text{pretest score}}{\text{max score} - \text{pretest score}}$$

The *N-gain* results were then used in descriptive and inferential analysis. The *N-gain* was calculated for each student and then averaged per class. Therefore, the reported mean *N-gain* is the average of individual *N-gain* scores, not a value calculated from the class mean *pretest* and *posttest* scores. The *N-gain* categories follow Hake (1998), namely high ($g \geq 0.7$), moderate ($0.3 \leq g < 0.7$), and low ($g < 0.3$). Before the t-

test, the *N-gain* data were tested for normality and homogeneity as prerequisites for parametric analysis.

a) Normality Test.

The normality test was conducted to determine whether the *N-gain* data in each group were normally distributed. The test used was the Lilliefors test (Kadir, 2015), with the hypotheses H_0 : the data are normally distributed and H_1 : the data are not normally distributed.

b) Homogeneity Test.

The homogeneity test was conducted to determine whether the variances of the *N-gain* of the experimental and control groups were homogeneous. The hypotheses tested were $H_0: \sigma_1^2 = \sigma_2^2$ and $H_1: \sigma_1^2 \neq \sigma_2^2$. Bartlett's chi-square test was used ($k = 2$ groups, $df = k - 1 = 1$):

$$X^2 = \ln(10) \left\{ B - \sum (n_i - 1) \log S_i^2 \right\}$$

where $B = \sum (n_i - 1) \log S_p^2$ and S_p^2 is the pooled variance. This test assumes independent observations and approximately normally distributed data within each group (which is supported here by the Lilliefors test), and it is sensitive to departures from normality. H_0 is rejected if the calculated value X^2 is greater than the critical value of X^2 ($\alpha = 0.05, df = 1$) of 3.841; otherwise, the variances are considered homogeneous.

c) Hypothesis Test.

The hypothesis test used an independent-samples t-test at the 5% significance level to determine the difference in the mean *N-gain* between the experimental group and the control group. The hypotheses tested were H_0 : there is no difference in the mean *N-gain* between the two groups and H_1 : there is a difference in the mean *N-gain* between the two groups. If $p < 0.05$, H_0 is rejected; conversely, if $p \geq 0.05$, H_0 fails to be rejected.

d) Effect Size.

In addition to statistical significance, the magnitude of the difference in improvement between the experimental group and the control group was analyzed using an effect size. Cohen's d was calculated based on individual *N-gain* scores to measure the standardized difference between the two groups:

$$d = \frac{M_{exp} - M_{ctrl}}{SD_{pooled}}$$

where (M_{exp}) and (M_{ctrl}) are the mean *N-gain* scores of the experimental and control groups, respectively, and (SD_{pooled}) is the pooled standard deviation, calculated as:

$$SD_{pooled} = \sqrt{\frac{(n_{exp} - 1)SD_{exp}^2 + (n_{ctrl} - 1)SD_{ctrl}^2}{n_{exp} + n_{ctrl} - 2}}$$

where (SD_{exp}^2) and (SD_{ctrl}^2) are the standard deviations of the individual *N-gain* scores, with $n_{exp} = 22$ and $n_{ctrl} = 20$. No small-sample correction was applied because Cohen's d was used as the effect size measure. According to Cohen (1988), values of 0.2, 0.5, and 0.8 indicate small, medium, and large effects, respectively. Given the quasi-experimental design without random assignment of individual students, the effect size represents the magnitude of the observed difference in *N-gain* between the two groups. This finding indicates that the experimental group obtained a greater learning

improvement than the control group after the implementation of the MAPACCING module.

3. Results

3.1 Results of Developing the MAPACCING Module

The product developed is the MAPACCING Learning Module, a Grade XI (Phase F) physics module on fluids based on STEM–*Design Thinking* and set in the context of coastal water quality in Majene. The module consists of usage instructions, a concept map, six learning activities (KB1–KB6), and a closing chapter with a summative test and a final reflection. It presents a contextual problem, namely the deteriorating quality of coastal well water, which is carried through all six activities. Each learning activity contains objectives, concept exploration, activities, a summary, comprehension exercises, and a reflection journal.

The activities follow the *Design Thinking* stages: Empathize, Define, Ideate, Prototype, Test, and Communication. The Science component covers fluid concepts, namely density, hydrostatic pressure, particle settling, and flow through porous media. Technology covers the principles of filtration, the characteristics of local materials, and the design criteria for the filter. Engineering is applied through designing alternative filters, building a prototype, and improving it based on test results. Mathematics is applied through calculating density, pressure, and discharge, and through analyzing data before and after filtration. Local materials, namely coconut husk, coconut shell charcoal, and sand, supported by gravel, *ijuk*, and filter cloth, are used in the design and construction of the filter. The coastal water quality problem is therefore not merely an illustration but the main context for the learning activities and the design of solutions.

Product development was carried out through stages that included the design process, expert assessment, revision, and product trial. At the development stage, expert assessment was used to obtain input on the product design before proceeding to the trial. This process is an important part of the 4D development model, because the results of the expert assessment are used as the basis for revising and refining the product before it is tried out with users (Sihombing et al., 2024). Table 3 summarizes the validators' comments, the revisions made, and the reasons for the revisions.

Table 3. Validator Comments, Revisions Made, and Reasons for Revisions

Section	Validator Comments	Revisions Made	Reason
Module cover (Figure 1)	The visual elements on the cover do not match the characteristics of the module and the physics material.	The visual elements of the cover were adjusted to the theme of physics and coastal water quality; less relevant elements were replaced.	The cover needs to reflect the identity and learning objectives of the module from the user's first impression.
Module content narrative (Figure 2)	The narrative section is too text-heavy.	Relevant illustrations were added to the narrative section.	Illustrations provide a concrete picture of the water quality issue, so readability improves.

Module format and instructions	Tables and the font type and size are not yet consistent; the instructions are not clear enough.	The table and font formats were standardized; the activity instructions were clarified.	A consistent appearance and clear instructions make it easier for students and teachers to use the module independently.
Activities in the module	The activities need to be aligned with the problem-solving skills indicators.	The activities were adjusted to the problem-solving skills indicators.	Problem-solving skills are the main objective of the module, so the activities must train them directly.
Practitioner response questionnaire	The practitioners' assessment criteria need to use a scale from very poor to very good.	The practitioners' assessment scale was arranged from very poor to very good.	A uniform scale makes it easier to categorize scores and interpret the results.
Student response questionnaire (instructions)	The questionnaire instructions do not need to be included in the student response sheet (this aspect obtained the lowest V, 0.56).	The questionnaire instructions aspect was removed from the student response sheet.	Students assess the module, not the questionnaire instructions; removing it keeps the items focused on assessing the module.
Student response questionnaire (content)	An aspect related to problem solving is needed.	Items on problem solving were added.	In line with the objectives of the MAPACING module, so that student responses also capture this aspect.
Problem-solving skills test	The presentation of data, the answer key, and the punctuation need to be improved.	The presentation of data in the questions, the answer key, and the punctuation were revised.	Clear questions and an accurate answer key reduce ambiguity in answers and maintain scoring consistency.

Source: Processed primary data, 2026

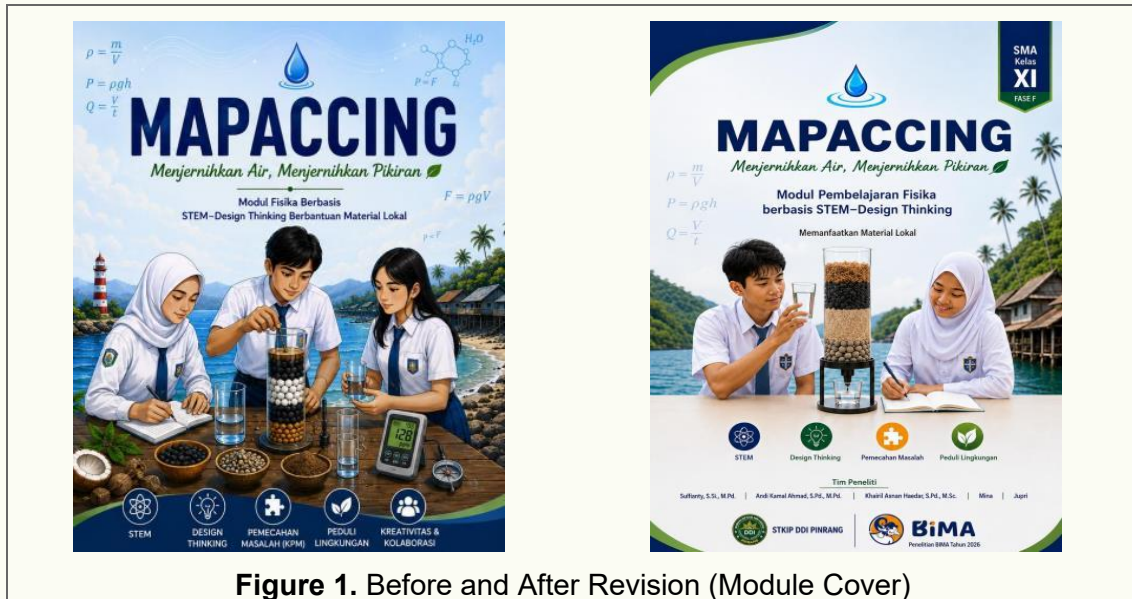


Figure 1. Before and After Revision (Module Cover)

Figure 1 shows the changes to the cover of the MAPACCING Module before and after revision. The validators judged that the visual elements on the cover did not match the characteristics of the module and the physics material; the revision was carried out by replacing and adjusting these elements so that the cover reflects the theme of physics and coastal water quality, which makes the identity and learning objectives of the module more visible from the first page.

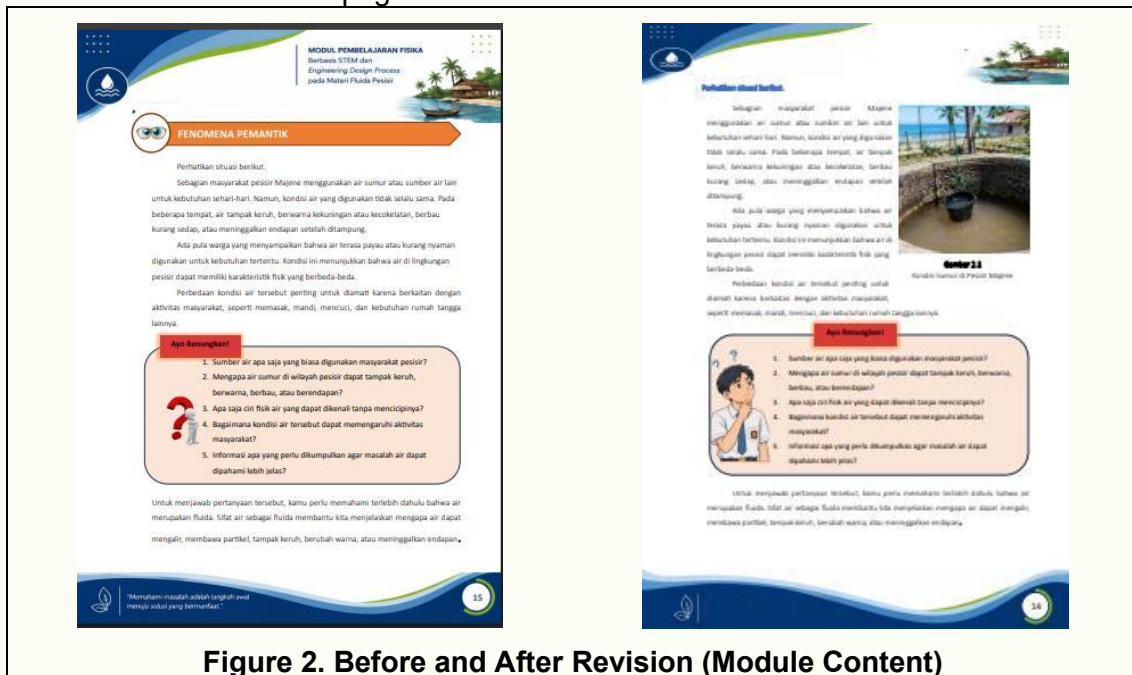


Figure 2. Before and After Revision (Module Content)

Revisions were also made to the content of the module (Figure 2). The validators judged that the narrative was too text-heavy; relevant illustrations were then added so that students would gain a concrete picture of the issue discussed. In addition, the consistency of tables, font type and size, and the clarity of instructions were improved, and the activities were aligned with the problem-solving skills indicators to match the objectives of the module. A summary of all revisions, including revisions to the questionnaires and the test, is presented in Table 3.

3.2 Results of the Expert Validation Analysis

Expert content validity is intended to determine the content validity of the developed product and the research instruments. After the instrument items were validated by experts, the scores obtained were calculated and then revised as necessary in accordance with the validators' suggestions.

Table 4. Validation Results of the MAPACING Module Using Aiken's V Index

Aspect Assessed	Total Item Validity Score	V	Category
Content Feasibility	16.56	0.79	Moderate
Presentation Feasibility	12.44	0.78	Moderate
Language Feasibility	13.67	0.72	Moderate
Graphics Feasibility	7.78	0.78	Moderate

Source: Processed primary data, 2026

Based on Table 4, all aspects of the MAPACING Module obtained Aiken's V values in the range of 0.72–0.79 and fell in the moderate category (0.40–0.80) according to Retnawati (2016), meeting the minimum practical criterion of $V \geq 0.40$, and were therefore considered acceptable (valid). The highest value was obtained for the content feasibility aspect at 0.79, while the lowest was for the language feasibility aspect at 0.72. These results indicate that the validators gave relatively consistent assessments of the suitability of the content, presentation, language, and graphics of the module.

Based on these validation results, the module could proceed to the trial stage after revisions were made based on the validators' input.

Table 5. Expert Validation Results for the Practitioner Response Questionnaire Using Aiken's V Index

Aspect Assessed	Total Item Validity Score	V	Category
Questionnaire Instructions	5.00	0.83	High
Content Feasibility	10.44	0.87	High
Presentation Feasibility	5.78	0.83	High
Language Feasibility	5.22	0.87	High
Graphics Feasibility	7.11	0.79	Moderate

Source: Processed primary data, 2026

All aspects of the practitioner response questionnaire obtained Aiken's V values in the range of 0.79–0.87 and fell in the moderate to high category (Retnawati, 2016), meeting the minimum practical criterion of $V \geq 0.40$, and were therefore considered acceptable (valid). The experts' assessment of the practitioner response questionnaire reached the same conclusion, namely that the questionnaire could be used with minor revisions.

Table 6. Content Validity Analysis of the Student Response Questionnaire Using Aiken's V Index

Aspect Assessed	Total Item Validity Score	V	Category
Questionnaire Instructions	3.33	0.56	Moderate
Content Feasibility	7.67	0.77	Moderate
Presentation Feasibility	6.33	0.90	High
Language Feasibility	3.67	0.73	Moderate
Graphics Feasibility	9.67	0.81	High

Source: Processed primary data, 2026

All aspects of the student response questionnaire obtained Aiken's V values of 0.56–0.90 and fell in the moderate to high category (Retnawati, 2016), meeting the minimum practical criterion of $V \geq 0.40$, and were therefore considered acceptable (valid). The highest value was found for the presentation feasibility aspect (0.90), while the lowest was for the questionnaire instructions (0.56). The experts concluded that the questionnaire could be used with minor revisions, including using a scale from very poor to very good for the students' assessment criteria, removing the questionnaire instructions from the student response sheet, and adding items related to problem solving in the content feasibility dimension, because this is an objective of the MAPACING module.

3.3 Analysis of the Problem-Solving Skills Test

The problem-solving skills test instrument initially consisted of 20 essay questions and was validated by three experts. Based on the results of the expert analysis, a mean Aiken's V of 0.84 was obtained, in the high category (Retnawati, 2016), so the instrument was declared suitable for use in the field with minor revisions related to the presentation of data, the answer key, and the clarity of punctuation.

After revision, the 20 test items were tested empirically in classes XII Physics 1 and Physics 2, which had already studied static and dynamic fluids. Based on item validity analysis, 12 valid items were obtained, with product-moment correlation coefficients greater than the r table value of 0.294 at the 0.05 significance level. The reliability analysis yielded a coefficient of 0.72 (reliable), which is theoretically considered reliable if the coefficient is at least 0.6 (Sugiyono, 2019). Thus, the 12 valid items were used as the problem-solving skills test instrument in the field test.

3.4 Practicality of the Module

a) Practicality of the Module Based on Student Responses

Table 7. Student Responses to the MAPACING Module

Class Interval	Frequency	Percentage (%)	Category
128–150	5	22.7	Very good
105–127	15	68.2	Good
82–104	2	9.1	Fair
59–81	0	0	Poor
36–58	0	0	Very poor
Total	22	100	

Source: Processed primary data, 2026

The analysis results show that 15 students (68.2%) responded in the good category, 5 students (22.7%) in the very good category, and 2 students (9.1%) in the fair category, with no students in the poor or very poor categories. The mean student response score was 116.6, which falls in the class interval of 105–127 with the criterion of good. This finding indicates that the MAPACING Module was well received by students as teaching material.

b) Practicality of the Module Based on Practitioner Responses

Table 8. Practitioner Responses to the MAPACING Module

Class Interval	Frequency	Percentage (%)	Category
117–137	1	33.33	Very good
96–116	2	66.67	Good
75–95	0	0	Fair
54–74	0	0	Poor
33–53	0	0	Very poor
Total	3	100	

Source: Processed primary data, 2026

One practitioner (33.33%) responded in the very good category, and two practitioners (66.67%) in the good category, with a mean score of 120.67 (class interval 117–137, very good category).

3.5 Results of Initial Effectiveness

a) Descriptive Statistics

The effectiveness of the MAPACING Module was analyzed based on changes in students' problem-solving skills (PSS) before and after instruction in the experimental class (using the MAPACING Module) and the control class (without the module). The PSS test consisted of 12 items with an ideal maximum score of 36.

Table 9. Descriptive Statistics of Problem-Solving Skills

Statistic	Experimental <i>Pretest</i>	Experimental <i>Posttest</i>	Control <i>Pretest</i>	Control <i>Posttest</i>
Sample Size (N)	22	22	20	20
Ideal Maximum Score	36	36	36	36
Ideal Minimum Score	0	0	0	0
Empirical Maximum Score	17	28	17	25
Empirical Minimum Score	6	15	5	13
Empirical Score Range	11	13	12	12
Empirical Mean Score	11.45	21.55	11.00	18.05
Standard Deviation	3.751	4.137	3.309	4.019
Variance	14.069	17.117	10.947	16.155

Source: Primary research data, 2026

Table 9 shows that the mean score of the experimental class increased from 11.45 on the *pretest* to 21.55 on the *posttest*, while the control class increased from 11.00 to 18.05. The increase in the mean score of the experimental class was 10.10 points, while

that of the control class was 7.05 points. Both classes experienced an improvement in problem-solving skills after instruction, but the improvement in the experimental class was greater than that in the control class.

The initial problem-solving skills of the two classes were examined using a standardized mean difference. The experimental class had a mean pretest score of 11.45 (SD = 3.75), while the control class had a mean pretest score of 11.00 (SD = 3.29), a mean difference of 0.45 points. The standardized mean difference was small (Cohen's $d = 0.13$; pooled SD = 3.55), indicating a small difference in initial problem-solving skills between the two classes. Although the two intact classes were randomly assigned to the experimental and control conditions, students were not individually randomized. Therefore, potential class-level differences should still be considered when interpreting subsequent differences in problem-solving gains.

b) *N-gain* Analysis

Table 10. *N-gain* Calculation Results for Problem-Solving Skills

Class	Mean Pretest	Mean Posttest	<i>N-gain</i>	Category
Experimental	11.45	21.55	0.42	Moderate
Control	11.00	18.05	0.29	Low

The mean *N-gain* of the experimental class was 0.42 (moderate category), while that of the control class was 0.29 (low category), which indicates that the improvement in problem-solving skills in the class using the MAPACING Module was higher than in the control class. The mean *N-gain* of the control class (0.29) is just below the 0.30 threshold, so the low category should be interpreted with caution, because this class is close to the boundary of the moderate category. The difference between groups is better judged from the t-test and the effect size than from the category labels.

c) *N-gain* Normality Test

Table 11. Normality Test Results for the Experimental and Control Classes

Data	Experimental	Control	Decision
N	22	20	Data are normally distributed
L Calculated	0.100	0.145	
L Critical	0.183	0.190	

Source: Processed primary data, 2026

The calculated L value of the *N-gain* was 0.100 for the experimental class and 0.145 for the control class, both smaller than the respective critical L values (0.183 and 0.190), so H_0 fails to be rejected and the *N-gain* data in both classes can be considered normally distributed.

d) *N-gain* Homogeneity Test.**Table 12.** Bartlett Homogeneity Test Results for the Variance of Problem-Solving Skills in the Experimental and Control Classes

Data	Experimental	Control
Variance	0.0074	0.0067
X ² Calculated (Bartlett)	0.05	
X ² Critical ($\alpha = 0.05, df = 1$)	3.841	

Source: Processed primary data, 2026

Bartlett's test yielded a calculated X² value of 0.05 (df = 1), which is smaller than the critical X² value of 3.841, so H₀ fails to be rejected and the variances of the *N-gain* of problem-solving skills in the experimental and control classes are considered homogeneous.

d) Hypothesis Test.

Based on the calculation results, the t table value for df = 40 and a significance level of 0.05 is 2.02, while the calculated t value (t₀) is 5.09, with a p-value of 0.000008775 (p < 0.001). Because the calculated t₀ > the t table value and p < 0.05, H₀ is rejected. Thus, there is a significant difference between the mean *N-gain* of the experimental class and that of the control class at the 5% significance level. The magnitude of this difference is indicated by the Cohen's *d* effect size = 1.57 (large category). In the context of this study, this result indicates initial evidence of the effectiveness of the MAPACCING Module in improving students' problem-solving skills.

4. Discussion

The results show that the MAPACCING Module, developed through the 4D model, is valid. Aiken's V values for the aspects of content feasibility, presentation, language, and graphics were in the range of 0.72–0.79, which falls in the moderate category (Retnawati, 2016). These results indicate that the components of the module were judged to be relevant and appropriate by the validators after improvements were made based on the suggestions given. Revisions were made mainly to the cover appearance, the density of the narrative text, format consistency, the clarity of instructions, the alignment of activities with problem-solving skills, and the presentation of the instruments. In terms of practicality, the mean student response of 116.6 was in the good category, while the mean practitioner response of 120.67 was in the very good category. These findings indicate that the module received positive responses from students, who used it in learning, and from practitioners, who observed the implementation of learning using the module and subsequently assessed its practicality.

In terms of initial effectiveness, both groups experienced an improvement in problem-solving skills after instruction. The mean score of the experimental class increased from 11.45 to 21.55, while the control class increased from 11.00 to 18.05. The mean *N-gain* of the experimental class of 0.42 was in the moderate category, while that of the control class of 0.29 was in the low category. The difference in *N-gain* between the two groups was also significant based on the independent-samples t-test (t = 5.09; p < 0.001), with a Cohen's *d* effect size of 1.57. These findings indicate that the group using the MAPACCING Module experienced a higher improvement in problem-solving skills than the control group. Although the two intact classes were randomly assigned to the experimental and control conditions, students were not individually randomized.

Therefore, the findings are more appropriately interpreted as initial evidence of effectiveness within the context of the two intact classes and should not be treated as strong causal evidence at the individual-student level.

The difference in improvement is consistent with the characteristics of the learning in the MAPACING Module, which integrates STEM–*Design Thinking*. The learning activities guide students to understand the water quality problem, identify needs, develop alternative solutions, design and build a filter prototype, test the results, and make improvements. This process gives students the opportunity to apply physics concepts in concrete problem-solving situations. The use of materials such as coconut husk, coconut shell charcoal, sand, and gravel also requires students to consider the function and arrangement of the materials in the filter design. Thus, the activities in the module give students room to understand the problem, plan solutions, carry out testing, and evaluate the results. This mechanism is an interpretation based on the characteristics of the learning activities and was not measured separately in this study.

The findings of this study are consistent with the study by Nguyễn et al. (2025), which showed that integrating *Design Thinking* into STEM learning can support the problem-solving process through repeated (iterative) activities. These results are also consistent with Syukri et al. (2018), who showed that the engineering design process in learning can support science problem-solving skills through activities of designing, making, testing, and improving. Research on STEM and the design process in physics learning also shows that design activities can give students opportunities to apply concepts in solving problems (Purwaningsih et al., 2020; Syukri et al., 2026; Winarno et al., 2020). The difference in this study lies in the integration of coastal water quality problems and local materials into STEM–*Design Thinking* activities.

The local context is an important part of the design of the MAPACING Module. The coastal water quality problem is used as the main learning context, while coconut husk, coconut shell charcoal, sand, and gravel are used in the activities of designing solutions. Thus, local materials are not merely practical materials but part of the problem that students need to consider when designing and testing solutions. This is consistent with the idea of *place-based education*, which emphasizes the connection of learning to students' local environment and experiences (Yemini et al., 2025). Nevertheless, this study did not specifically measure students' motivation, engagement, or perceptions of the meaningfulness of the local context. Therefore, the contribution of the local context to the improvement in problem-solving skills cannot yet be empirically separated from the contribution of the STEM–*Design Thinking* approach.

Not all findings showed optimal results. The *N-gain* of the experimental class of 0.42 was still in the moderate category, so the improvement in problem-solving skills was not yet maximal for all students. The highest *posttest* score also reached only 28 out of a maximum score of 36. In addition, the control class still experienced an *N-gain* improvement of 0.29. The improvement in the control group shows that learning without the MAPACING Module can also improve problem-solving skills, so changes in students' scores are not related only to the use of the module. In terms of the product, the language aspect obtained the lowest validation value, namely 0.72, and two students responded in the fair category. The test reliability of 0.72 also indicates that the instrument can still be developed further. These findings indicate that there is still room to improve the readability of the module, strengthen the instruments, extend the learning time, and provide more varied problem-solving experiences.

In terms of implications, the MAPACING Module can be used as an alternative physics teaching material that connects physics concepts with environmental problems through STEM–*Design Thinking* activities and local materials. The results of this study

also provide a basis for further development involving more schools and students, longer learning time, and stronger measurement instruments. Because this study was conducted in one school with a relatively limited number of students, the results need to be viewed as initial effectiveness that still requires testing in broader contexts.

5. Conclusion

This study produced the MAPACCING Physics Learning Module based on STEM–Design Thinking with local materials in the context of coastal water quality, developed through the 4D model. Within the testing context of this study, expert validation showed that the module met the criteria for content validity, and student and practitioner responses indicated good practicality in its implementation.

The main evidence of initial effectiveness was shown by the greater improvement in problem-solving skills in the experimental class than in the control class, with a statistically significant difference and a large effect size. These findings provide initial empirical support for the use of the MAPACCING Module in supporting students' problem-solving skills within the context of the two intact classes studied. The contribution of this study is to provide a model of physics teaching material that integrates STEM, Design Thinking, local materials, and problem solving, accompanied by initial empirical evidence from the coastal context of Majene.

5.1 Research Limitations

The findings of this study need to be interpreted with caution because of several limitations: (1) the implementation was carried out in only one school; (2) the sample size was small (22 students in the experimental class and 20 students in the control class); (3) students were not randomly assigned because intact classes that had already been formed were used, and with only one class per condition, class-level differences cannot be separated from the effect of the module; (4) both classes were taught by the research team, so the potential for researcher bias cannot be fully ruled out; (5) the duration of the intervention was relatively short; (6) the separate contribution of the coastal context to problem-solving skills was not tested because the module was evaluated as a whole; and (7) the language validation was carried out by a validator with a background in English Language Education, while the module was written in Indonesian. Therefore, the result of the language feasibility assessment ($V = 0.72$), which is the lowest compared with the other aspects, needs to be interpreted by considering the validator's expertise background.

Thus, the conclusions regarding the validity, practicality, and initial effectiveness of the module apply within the testing context carried out. Generalizing the findings to other schools or populations needs to be done cautiously and requires further testing in more diverse contexts.

5.2 Recommendations

Based on the results of this study, the MAPACCING Module can be used as an alternative physics teaching material that integrates the STEM–*Design Thinking* approach with the context of coastal water quality and the use of local materials. The use of this module can be developed further for other topics and local environmental problem contexts to expand its application in physics learning.

Future research is recommended to involve more schools and students, as well as more diverse school characteristics, so that the effectiveness of the module can be tested more broadly. In addition, further research can be conducted to examine the effect

of using the MAPACCING Module over a longer period, and to examine in more depth the development of problem-solving skills at each stage of STEM–Design Thinking.

6. Acknowledgments

The authors would like to thank the Ministry of Higher Education, Science, and Technology for the funding support provided through the 2026 Beginner Lecturer Research (Penelitian Dosen Pemula, PDP) scheme, as well as the Institute for Research and Community Service (LPPM) of STKIP Darud Da'wah Wal Irsyad Pinrang and SMA Negeri 3 Majene for their support in carrying out this study.

References

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- Arikunto, S. (2010). *Prosedur penelitian: Suatu pendekatan praktik* [Research procedures: A practical approach]. Rineka Cipta.
- Awang, H., Abdullah, P. S., Abdul Latiff, Z. A., Huda Mohd Azmin, S. N., & Boukhliji, F. (2025). Coconut shell-derived biocarbon composite for adsorptive field-water purification. *Materials Chemistry and Physics*, 333, 130359. <https://doi.org/10.1016/j.matchemphys.2025.130359>
- Azwar, S. (2012). *Reliabilitas dan validitas* [Reliability and validity]. Pustaka Pelajar.
- Che Osmi, S. K., Suif, Z., & Aripin, A. M. E. (2024). Design and fabricate of coconut shell activated carbon for groundwater filtration system. *Advances in Agricultural and Food Research Journal*, 5(1). <https://doi.org/10.36877/aafri.a0000420>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- De Sá, I. A., Costa, A. L. S., Fernandes Santos, E. M., & De Oliveira, A. K. C. (2019). Study of efficiency of oil removal of water produced (Petrobras RN/CE) using adsorbent materials. *The Journal of Engineering and Exact Sciences*, 5(2), 203–206. <https://doi.org/10.18540/jcecvl5iss2pp0203-0206>
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Hasanah, R., Gunawan, G., Muntari, M., Rokhmat, J., & Harjono, A. (2025). Development of a problem-based physics teaching module to enhance students' science process skills. *Jurnal Pendidikan Fisika dan Teknologi*, 11(1), 257–271. <https://doi.org/10.29303/jpft.v11i1.8384>
- Kadir. (2015). *Statistika terapan: Konsep, contoh, dan analisis data dengan program SPSS/Lisrel dalam penelitian* [Applied statistics: Concepts, examples, and data analysis using SPSS/Lisrel in research]. PT RajaGrafindo Persada.
- Kania, N., Kusumah, Y. S., Dahlan, J. A., Nurlaelah, E., Gürbüz, F., & Bonyah, E. (2024). Constructing and providing content validity evidence through the Aiken's V index based on the experts' judgments of the instrument to measure mathematical problem-solving skills. *REID (Research and Evaluation in Education)*, 10(1), 64–79. <https://doi.org/10.21831/reid.v10i1.71032>

- Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. (2022). *Kurikulum Merdeka: Kerangka dasar* [Merdeka Curriculum: Basic framework].
- Khaeruddin, & Ali, M. S. (2012). *Evaluasi pembelajaran* [Learning evaluation]. Badan Penerbit UNM.
- Kusnandar, D. (2025). Pembelajaran IPA berbasis lingkungan pesisir untuk meningkatkan literasi sains siswa Madrasah Tsanawiyah [Coastal environment-based science learning to improve Madrasah Tsanawiyah students' scientific literacy]. *Multidisciplinary: Journal of Education and Learning*, 1(2), 71–75. <https://doi.org/10.68195/multidisciplinary.v1i2.13>
- Nguyễn, L. C., Hoa, H. Q., & Hien, L. H. P. (2025). Integrating design thinking into STEM education: Enhancing problem-solving skills of high school students. *Eurasia Journal of Mathematics, Science and Technology Education*, 21(4), em2611. <https://doi.org/10.29333/ejmste/16084>
- Pagiling, S. L., Munfarikhatin, A., Natsir, I., & Ma'dika, A. R. (2024). Development of mathematics teaching modules based on differentiated instruction to stimulate student engagement. *Jurnal Didaktik Matematika*, 11(1), 21–40. <https://doi.org/10.24815/jdm.v11i1.36525>
- Purwaningsih, E., Sari, S. P., Sari, A. M., & Suryadi, A. (2020). The effect of STEM-PjBL and discovery learning on improving students' problem-solving skills of the impulse and momentum topic. *Jurnal Pendidikan IPA Indonesia*, 9(4), 465–476. <https://doi.org/10.15294/jpii.v9i4.26432>
- Razzouk, R., & Shute, V. (2012). What is design thinking and why is it important? *Review of Educational Research*, 82(3), 330–348. <https://doi.org/10.3102/0034654312457429>
- Retnawati, H. (2016). *Validitas, reliabilitas, & karakteristik butir* [Validity, reliability, & item characteristics]. Parama Publishing.
- Sihombing, B., Zamsiswaya, & Sawaluddin. (2024). Model pengembangan 4D (Define, Design, Develop, dan Disseminate) dalam pembelajaran pendidikan Islam [The 4D development model (Define, Design, Develop, and Disseminate) in Islamic education learning]. *Journal of Islamic Education El Madani*, 4(1), 11–19. <https://doi.org/10.55438/jiee.v4i1.135>
- Sudjana. (1992). *Metode statistika* [Statistical methods] (5th ed.). Tarsito.
- Sugiyono. (2019). *Metode penelitian dan pengembangan* [Research and development methods]. Alfabeta.
- Sulfianty, S., Arsyad, M., & Helmi, H. (2024). Development of student worksheets for environmental based physics learning with a STEM approach for students SMA Negeri 1 Sendana. *International Journal of Social Science and Human Research*, 7(3), 2007–2014. <https://doi.org/10.47191/ijsshr/v7-i03-64>
- Sunarwan, B., & Firmansyah, D. (2016). Tinjauan kimia airtanah Kabupaten Majene, Sulawesi Barat [A review of the groundwater chemistry of Majene Regency, West Sulawesi]. *Jurnal Teknik: Majalah Ilmiah Fakultas Teknik UNPAK*, 17(1), 9–15. <https://doi.org/10.33751/teknik.v17i1.911>
- Syukri, M., Soewarno, S., Halim, L., & Mohtar, L. E. (2018). The impact of engineering design process in teaching and learning to enhance students' science problem-solving skills. *Jurnal Pendidikan IPA Indonesia*, 7(1), 66–75. <https://doi.org/10.15294/jpii.v7i1.12297>

- Syukri, M., Haya, F., Maghfirah, S., Herliana, F., Rasul, M. S., & Pratama, H. (2026). STEM-based engineering design process models in physics learning: Systematic literature review. *European Journal of STEM Education*, 11(1), Article 4. <https://doi.org/10.20897/ejsteme/17784>
- Utami, A., Astuti, D., Sandi, A. P., & Astuti, D. P. (2023). STEM-based module outline, validity, readability and response towards layout, content and utility aspect. *Journal of Science Education Research*, 7(1), 55–66. <https://doi.org/10.21831/jser.v7i1.51371>
- Wahyuni, S., Indrawati, I., Sudarti, S., & Suana, W. (2017). Developing science process skills and problem-solving abilities based on outdoor learning in junior high school. *Jurnal Pendidikan IPA Indonesia*, 6(1), 165–169. <https://doi.org/10.15294/jpii.v6i1.6849>
- Winarno, N., Rusdiana, D., Samsudin, A., Susilowati, E., Ahmad, N. J., & Afifah, R. M. A. (2020). Synthesizing results from empirical research on engineering design process in science education: A systematic literature review. *Eurasia Journal of Mathematics, Science and Technology Education*, 16(11), em1912. <https://doi.org/10.29333/ejmste/9129>
- Yassin, M., Abba, S., Shah, S., Usman, A., Egbueri, J., & Agbasi, J. (2024). Toward decontamination in coastal regions: Groundwater quality, fluoride, nitrate, and human health risk assessments within multi-aquifer Al-Hassa, Saudi Arabia. *Water*, 16(10), 1401. <https://doi.org/10.3390/w16101401>
- Yemini, M., Engel, L., & Ben Simon, A. (2025). Place-based education – A systematic review of literature. *Educational Review*, 77(2), 640–660. <https://doi.org/10.1080/00131911.2023.2177260>
- Zahra, L., Ghazali, M. F., Dermawan, A., Salsabila, C., Aulia, M., & S, N. M. M. (2022). Karakteristik intrusi air laut serta hubungannya dengan ketinggian dan kemiringan lereng [Characteristics of seawater intrusion and its relationship with elevation and slope]. *Jurnal Geologi Kelautan*, 20(2), 100–107. <https://doi.org/10.32693/jgk.20.2.2022.772>