PROFESI PENDIDIKAN DASAR
Volume 12, Issue 2, Month yyyy, pp. 182-198
p-ISSN: 2503-3697, e-ISSN: 2541-2590
https://journals2.ums.ac.id/index.php/ppd

Revealing Widespread Misconception: A Multitier Science Instrument (MscI) to Assess Pre-Service Elementary Teachers' Understanding

Dewi Yulianawati[1], Eliya Rochmah[1*], Widia Nur Jannah[1], Fikriyah[1], Fanny Septiany Rahayu[1]
[1]Elementary Teacher Education Study Program, Faculty of Teacher Training and Education, Universitas Muhammadiyah Cirebon, Indonesia
*Corresponding Author's email: eliya.rochmah@umc.ac.id

How to cite: Yulianawati, D., Rochmah, E., Jannah, W. N., Fikriyah, F., & Septiany Rahayu, F. (2025). Revealing Widespread Misconception: A Multitier Science Instrument (MscI) to Assess Pre-Service Elementary Teachers' Understanding. *Profesi Pendidikan Dasar*, *12*(2), 182–198. https://doi.org/10.23917/ppd.v12i2.10069

Keywords:
misconceptions; diagnostic instrument; multitier instrument; elementary teacher

Article History:
Submitted: 2025-06-01
Revised: 2025-08-01
Accepted: 2025-08-27

Abstract

Prior studies demonstrate that numerous pre-service teachers retain scientific misconceptions, and existing assessment instruments are inadequate for accurately distinguishing levels of comprehension. This research developed the Multitier Science Instrument (MScI), a reliable diagnostic tool for improving the evaluation of conceptual comprehension. The research utilised the 4D model, which consists of four phases: defining, designing, developing, and disseminating. The study involved thirty-one pre-service elementary school teachers, with an average age of 19 years, at an Islamic university in Cirebon, Indonesia. CVR analysis and other statistical techniques evaluated the instrument's validity and reliability while analysing the distribution of misconceptions. The results demonstrate that the Multitier Science Instrument (MScI) possesses high validity, evidenced by a Content Validity Ratio of 0.966, surpassing the benchmark of 0.672, and robust reliability, as indicated by a Cronbach's alpha coefficient of 0.827. The responses from participants were divided into four groups: sound understanding, partial understanding, misconceptions, and no understanding. Fluid statics exhibited the highest degree of misconceptions among the evaluated subjects. These findings underscore the pressing necessity for enhanced diagnostic instruments in education. Conventional assessment techniques, such as multiple-choice examinations and essay prompts, inadequately reflect the genuine depth of pre-service elementary educators' conceptual comprehension.

INTRODUCTION

Background of the Study

The 21st century has seen tremendous progress in information dissemination and technology, which has led to revolutionary improvements in a number of fields. Innovation in technology is greatly aided by science. In order to succeed in the workforce of the twenty-first century, people need to develop a strong foundation in science. Teachers are expected to demonstrate essential 21st-century competencies, particularly effective problem-solving abilities, within the framework of Society 5.0. A significant obstacle in science education is the students' struggle to resolve science-related problems. Their poor problem-solving skills stem from their incapacity to apply scientific concepts to the current problems (Yulianawati et al., 2018). This implies that students' ability to solve problems effectively is significantly impacted by their conceptual understanding.

© The Author(s). 2025
This work is licensed under a Creative Commons Attribution 4.0 International License

A solid grasp of scientific principles is an essential aspect of education, especially for prospective elementary teachers who will be the primary facilitators of knowledge for their future students. However, many studies have shown that misunderstandings about science are still a big problem at all levels of education, even at the university level (Aydeniz et al., 2017; Silva Mangiante, 2025). Pre-service elementary teachers with misconceptions in scientific content may inadvertently convey these inaccuracies to their students, potentially resulting in a long-term negative impact on the quality of elementary science education.

Individuals often develop varying interpretations of scientific concepts, and when these interpretations deviate from scientifically accepted explanations, they may evolve into misconceptions (Mufida et al., 2024). Several areas within physics have been widely recognised as particularly susceptible to such misconceptions, including work and energy (Afif et al., 2017), temperature, and heat (Safitri et al., 2025), static electricity (Hermita et al., 2017), and Newton's laws of motion (Fratiwi et al., 2017). Students' understandings are often influenced by their daily experiences and intuitive interpretations of physical phenomena (Turgut & Gurbuz, 2012; Yalçin et al., 2017). This indicates that learners do not arrive in the classroom as blank slates to be filled with information; instead, they come equipped with various prior knowledge and experiences. Duit and Treagust (2003) stress that conceptual understanding is shaped not only by formal education but also by learners' personal experiences and their unique interpretations of natural phenomena. Consequently, there exists an urgent necessity for enhanced diagnostic techniques to detect and evaluate misconceptions, particularly among pre-service elementary educators, in order to promote the cultivation of precise scientific understanding.

Problem of the Study

Pre-service elementary teachers are well aware of the persistent misconceptions in science (Aydeniz et al., 2017; Silva Mangiante, 2025). The principal concern is that these prospective educators may hold beliefs that conflict with scientifically accepted principles and may, often inadvertently, transmit them to their pupils. Such transmission is particularly troubling, as it can hinder the development of scientific understanding from the earliest stages of schooling (Al Sultan et al., 2018). Addressing this issue necessitates both effective pedagogical approaches and the use of valid, reliable diagnostic instruments capable of identifying misconceptions in sufficient depth. Nonetheless, numerous frequently utilised assessments—particularly in certain higher education contexts, such as Islamic universities in Cirebon—continue to be predominantly characterised by single-tier multiple-choice items and essay questions. These instruments are insufficient for examining the rationale behind students' responses and fail to measure students' confidence or conviction in their answers (Yan & Subramaniam, 2018).

Inconsistencies arise when traditional assessment tools fail to demonstrate how well aspiring teachers truly grasp concepts. Due to the constrictive nature of tests that only focus on final responses and fail to analyse students' thought processes, many misconceptions go unnoticed. According to recent research, single-tier multiple-choice exams are ineffective at revealing students' thought processes and additional ideas (Yeo et al., 2022). As a result, teachers' perceptions of their students' knowledge and their actual knowledge diverge. The underutilisation of multitier diagnostic tools in teacher education and elementary education frameworks reveals yet another knowledge gap. In order to better distinguish between accurate comprehension and misconceptions, a number of studies have created two-tier (Kamcharean & Wattanakasiwich, 2016), three-tier (Gurcay & Gulbas, 2015; Samsudin et al., 2017), and even four-tier tests (Afif et al., 2017; Fratiwi et al., 2017; Kaltakci-Gurel et al., 2017). However, the majority of the time, these tests are only utilised in secondary and higher education or research settings. Promoting sound scientific understanding requires early detection of misconceptions among aspiring educators.

Research's State of the Art

Concepts are described by Özkan and Selçuk (2012) as representations of the general characteristics of similar things and events that can be expressed using a variety of words . Personal conceptions are the various ideas that people may have about a particular concept. Students have a variety of preconceptions prior to formal education that are shaped by their everyday experiences and interactions with their environment. According to Suhandi et al. (2020), students' conceptions are internal cognitive constructs that are based on external representations provided by people like teachers, authors of textbooks, or designers of digital media.

According to Özkan and Selçuk (2015), preconceptions encompass all of the beliefs that students hold prior to learning and are crucial because they significantly impact students' learning outcomes. Students must actively participate in expanding their own knowledge if they are to learn effectively (Lombardi et al., 2021). When it comes to developing scientific understanding, preconceptions are a good place to start (Menia et al., 2017). In order to better plan and deliver instruction, teachers must ascertain students' preconceived ideas about the scientific concepts they will be teaching.

Some assumptions can be helpful for learning, but others might go against what scientists already know. Misunderstandings occur when students' ideas diverge from scientifically accepted theories or the consensus of experts (Suprapto, 2020). They do not arise by chance; rather, they originate from identifiable sources. Galanis et al. (2025) and Kaltakci and Didis (2007) assert that the primary sources of misconceptions are students' personal experiences, the instructional language, textbooks, and, most importantly, the teacher. Guerra-Reyes et al. (2024) provide additional detail by grouping the reasons for misconceptions into three main categories. First, misunderstandings may arise when students interpret new classroom information through the prism of their everyday experiences. Second, they may originate from inaccurate or incomplete prior knowledge, which hinders the development of accurate scientific understanding. Third, when authoritative sources, such as textbooks or teachers, present information incorrectly, this can also lead to misunderstandings. In short, misconceptions are ideas that students hold that are not in line with what scientists believe. They come from a variety of places that students encounter in school.

A variety of diagnostic tools have been created to find misconceptions in science education. These tools usually come with systematic steps that are meant to bring out students' conceptual misunderstandings. Open-ended tests, structured interviews, and multiple-choice assessments are all common ways to do this (Maryati & Priatna, 2018; Ilyas & Saeed, 2018; Eshach et al., 2018). According to Samsudin et al. (2024), open-ended tests and interviews require considerable time for evaluation. Consequently, multiple-choice tests are often preferred for their relative efficiency and reduced susceptibility to bias, enabling consistent evaluation of responses. Over recent decades, substantial attention has been devoted to refining multiple-choice tests through the development of multi-tier formats, including two-tier (Kamcharean & Wattanakasiwich, 2016), three-tier (Gurcay & Gulbas, 2015; Samsudin et al., 2017), and four-tier tests (Afif et al., 2017; Fratiwi et al., 2018; Kaltakci-Gurel et al., 2017). These new ideas get around the problems with traditional formats by giving us more information about how students think and how sure they are of themselves. Samsudin (2024) pointed out that the four-tier test has been very good at telling the difference between scientifically correct ideas and wrong ones.

A two-tier diagnostic test has two different levels for each item. Laliyo et al. (2019) stress the importance of having an open-ended part in the second tier so that students can explain their reasoning when none of the choices fit with what they know. This design makes it less likely that students will guess, which gives teachers a better idea of what students are getting wrong and how bad it is. The three-tier test builds upon the two-tier format by incorporating an additional assessment layer, requiring students to indicate their confidence in their chosen answers (Gurcay & Gulbas, 2015; Samsudin et al., 2017). This structure enhances its utility in diagnosing misconceptions, evaluating the reasoning underpinning responses, and distinguishing between a lack of knowledge and a misconception. The articulated confidence level functions as a diagnostic metric, providing significant insights into students' assurance and comprehension depth. The four-tier test represents a further refinement of this approach. The first tier requires selecting an answer to a specific question; the second tier assesses confidence in the selected answer; the third tier evaluates the reasoning behind the choice; and the fourth tier gauges confidence in that reasoning (Kaltakci-Gurel et al., 2017). According to Samsudin (2024), the four-tier test has proven highly effective in differentiating scientifically accurate conceptions from misconceptions.

Gap Study and Objective

Misconceptions in science education have been a central focus of numerous studies, owing to their substantial impact on students' conceptual understanding of scientific content. Different diagnostic tools, like two-tier tests (Kamcharean & Wattanakasiwich, 2016), three-tier (Gurcay & Gulbas, 2015; Samsudin et al., 2017), and four-tier tests (Afif et al., 2017; Fratiwi et al., 2018; Kaltakci-Gurel et al., 2017), have been made and shown to be better at finding misconceptions than regular single-tier tests (Hermita et al., 2017; Samsudin et al., 2024). Nonetheless, the utilisation of these multitier instruments is still restricted and infrequently employed in teacher education, especially for pre-service elementary educators. This gap is of critical concern, as the early detection of misconceptions during teacher preparation is essential to prevent their transmission to future pupils and to foster the development of accurate scientific literacy from the very outset of schooling.

Moreover, the majority of existing diagnostic instruments focus on a single mode of representation, thereby limiting their capacity to provide a comprehensive account of students' conceptual understanding. Previous research has shown that combining different types of representation, such as verbal, visual, graphical, and symbolic, can help us better understand how students think and how they understand concepts (Aminuddin et al., 2024; Pramonoadi et al., 2020). Many teacher education programs use essay-based tests and single-tier multiple-choice tests, which do not really test students' thinking skills or how sure they are of their answers. When teacher education programs rely too much on narrowly defined assessment formats, it makes it harder to get deep diagnostic insights and come up with targeted interventions to fix specific misunderstandings. This makes science instruction less effective overall.

This demonstrates a clear discrepancy between the practical approaches currently used in teacher education and theoretical developments in the assessment of misconceptions. The lack of a comprehensive, valid and multirepresentational multitier instrument, particularly for evaluating misconceptions about static fluid among pre-service elementary teachers, represents a significant gap warranting further investigation. In response to these issues, this study seeks to address the gap by developing the Multitier Science Instrument (MScI), which integrates multiple forms of representation within each item. The MScI is specifically designed to identify and analyse misconceptions that pre-service elementary teachers hold regarding science concepts such as static fluid. It has been designed as a more effective and dependable substitute for traditional tests, offering a more comprehensive and precise assessment of conceptual comprehension. The results obtained from its implementation are anticipated to guide the development of more effective learning strategies based on a comprehensive understanding of students' actual conceptual comprehension.

METHOD

Type and Design

Figure 1 illustrates the use of 4D in this research design, which stands for Define, Design, Develop, and Disseminate.

Data and Data Sources

This study included 31 pre-service elementary teachers in their second semester, with a mean age of 19 years (7 men and 24 women). They were students at an Islamic university in Cirebon, Indonesia, which is on the north coast of West Java Province in the eastern part of the country. The sample was chosen using purposive sampling, and the only requirement for participants was that they were currently taking the Science Concepts course

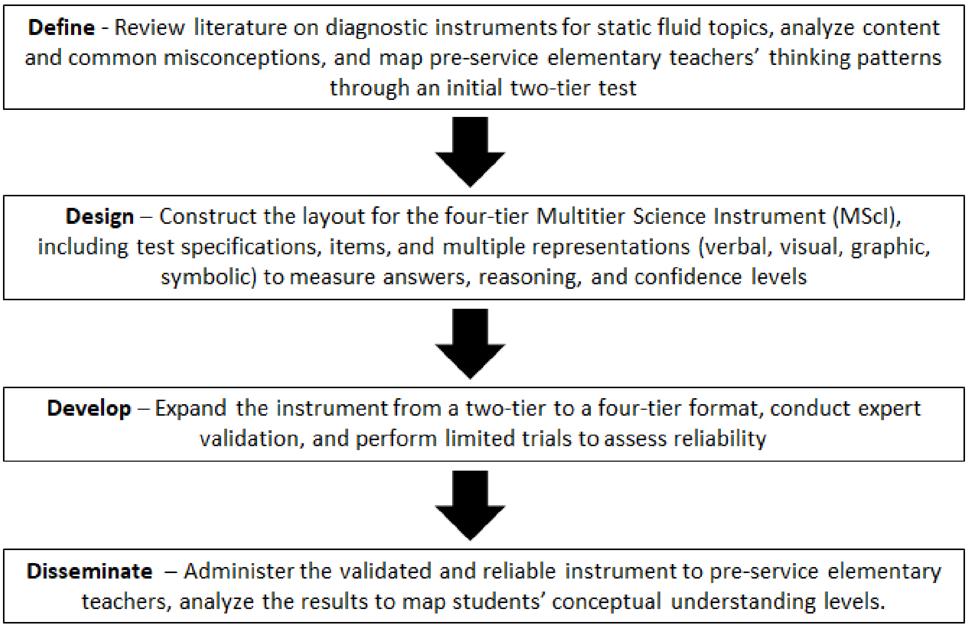


Figure 1. 4D Design Scheme

Data Collection Technique

Data collection procedures were designed to evaluate the validity, reliability and conceptual proficiency of pre-service elementary teachers, as summarised in Table 1. To determine validity, six subject-matter experts appraised each MScI item utilising a structured validation sheet. This sheet included indicators like how easy it was to read the questions, how well the concepts fit, how reasonable the content was, how well the third-tier reasoning matched the first-tier answer choices, and how well it worked for finding certain misconceptions. The Content Validity Ratio (CVR) was then calculated to determine item validity.

Following validation, the MScI instrument was administered to the pre-service elementary teachers. The total score of each participant was utilised to evaluate the instrument's reliability, while the responses at each level were examined to determine the teachers' conceptual comprehension of each MScI item.

Table 1. Data Collection Instruments and Techniques

Data	Research Instrument	Data Collection Techniques
Validity of MScI	Validation Questionnnaire	Judgement/Expert Review
Reliabity of MScI	MScI	Test
Pre-service elementary teachers' level of conception	MScI	Test

Data Analysis

Three types of data were gathered during this study's dissemination phase: individual student responses at each tier for the codification of conceptual understanding levels, aggregate student scores for the reliability analysis, and expert validation to ascertain the MScI instrument's content validity. A thorough assessment of the instrument's psychometric qualities was made possible by these data sources. The Content Validity Ratio (CVR) was used to measure the level of expert consensus. The CVR represents a linear transformation of the proportion of experts who rated an item as "valid without revision" and is computed as follows:

$$CVR = \frac{n_e - \left(\frac{N}{2}\right)}{\left(\frac{N}{2}\right)}$$

In this context, CVR denotes the Content Validity Ratio, n_e is the number of experts who said the item was "valid without revision," and N represents the total number of experts participating in the validation process. A CVR value of 1 indicates unanimous expert agreement on the item’s importance. When exactly half of the experts rate the item as important, the CVR value falls between 0 and 1. On the other hand, the CVR value becomes negative if fewer than half of the experts think the item is important. Wilson et al. (2012) say that the lowest CVR score that six experts will accept is 0.672.

The test–retest method was used to check how reliable the MScI instrument was. This method involves giving the same instrument to the same group of people twice. The Pearson product–moment correlation formula was used to find the reliability coefficient, and the scores were interpreted using the rules set out by Guilford (1956) (Guilford, 1956). A high correlation value means that the tool gives the same results over time, which shows that it is stable and reliable as a measurement tool.

Responses provided by students to each tier of the MScI were used to generate item-specific codes, with symbols assigned according to the criteria presented in Table 2. These codes indicate the level of conception demonstrated for each item. Then, scores were given based on the level of conception that was found. Sound Understanding, Partial Understanding, Misconception, No Understanding, and Uncodable are the different levels of conception (Samsudin et al., 2016) . This systematic coding allows for precise categorisation of students' conceptual comprehension and fosters uniformity in the interpretation of assessment outcomes.

Table 2. MScl Answer Criteria

Conception Level	**Symbol**	***Tier - 1***	***Tier – 2***	***Tier -3***	***Tier - 4***	**Score**
Sound Understanding		T	S	T	S	2
Partial Understanding		T	S	T	NS	1
		T	NS	T	S	
		T	S	T	NS	
		T	NS	F	S	
		T	S	F	NS	
		T	NS	F	S	
		T	NS	F	NS	
		F	S	T	S	
		F	S	T	NS	
		F	NS	T	S	
		F	NS	T	NS	
Misconception		F	S	F	S	0
No Understanding		F	S	F	NS	0
		F	NS	F	S	
		F	NS	F	NS	
Uncodable		Incomplete responses on the test instrument				0

Notes: T = True, F = False, S = Sure, and NS = Not Sure

RESULTS

Define

In the Define stage, the activities undertaken included the analysis of various diagnostic test instruments, specifically the two-tier, three-tier, and four-tier formats. The four-tier diagnostic test is the most accurate way to see how well students understand concepts, according to many people. The first tier consists of multiple-choice answers; the second tier measures how sure the person is about their answer; the third tier asks for an explanation of why they chose that answer; and the fourth tier checks how sure the person is about their explanation. Following this, an analysis was undertaken of the static fluid topic, which encompassed hydrostatic pressure, buoyant force, and the positioning of objects in a liquid, within the context of pre-service elementary teacher education.

The researcher also used a two-tier diagnostic tool to find out how well students understood static fluid concepts. In the first tier, students chose an answer, while in the second tier, they provided a justification for their chosen response based on their conceptual understanding. This structure was purposefully crafted to ensure that the reasoning options at the third tier of the MScI instrument could subsequently identify potential patterns of student cognition, encompassing scientifically accurate conceptions, misconceptions, and partial understanding.

Design

At the design stage, the MScI was developed using multiple forms of representation, including images, verbal explanations, diagrams, symbols and graphs. Table 3 shows how the MScI questions are spread out. The instrument was organised into four levels, as shown in Figure 2, and is described as follows:

a. The first level has five answer choices, one of which is correct and four of which are wrong.
b. The second tier required students to indicate their level of confidence in the answers provided in the first tier. The available options for this tier were limited to "Yes" and "No.
c. The third tier was presented in the form of semi-open-ended questions. Answer choices A, B, C and D represented reasoning options identified during the design-stage analysis, while option E provided an open field for students to articulate alternative reasoning not captured in the predefined options. The reasoning statements in options A to D were not invariably correct; therefore, the most appropriate reasoning might, in some cases, be expressed by students in option E.
d. The fourth level asks how sure you are about the reasoning in the third level, and you can choose between "Yes" and "No" levels of confidence.

Question
Options A. B. C. D. E.
Are you confident in your answer? A. Yes B. No
Reason A. B. C. D. E. ....
Are you confident of the reason you gave? A. Yes B. No

Figure 2. MScI Design

Table 3. Distribution of MScI

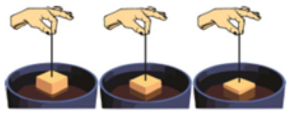
Sub Topic	Question Item Number	Type of Representation	
		Description	Options
Hysrostatic Pressure	Q1	P	V
	Q2	P	S
	Q3	P	D
	Q4	P	V
Buoyant Force	Q5	V	V
	Q6	P	S
	Q7	D	S
	Q8	P	G
	Q13	P	D
Position of objects in liquid	Q9	P	P
	Q10	P	V
	Q11	P	V
	Q12	P	V

Note:
P: Picture; V:Verbal; D: Diagram ; S: Symbol; G: Graph

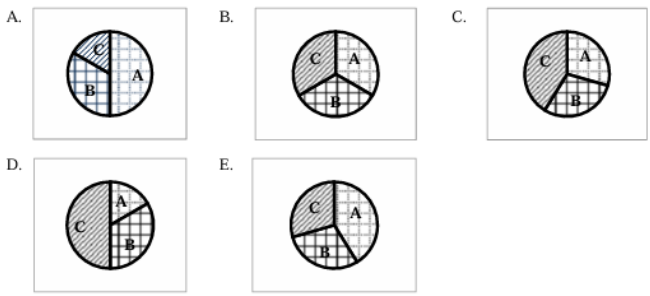
Develop

During the Develop stage, thirteen MScI items were created (see Figure 3; https://bit.ly/45Bcipz). At this point, the researchers added the parts of the question, the answer choices for the first tier, and the reasoning choices for the third tier. The finished MScI instrument, as shown in Figure 3, was the result of this stage.

13.1. Three containers (A, B, and C) are filled with the same liquid and with the same volume. Each container holds an object that is identical material, mass, and volume. The three objects are placed into the liquid as shown in the picture below.



Based on the picture above, which of the following diagrams correctly shows the agnitude of the buoyant force acting on the objects in containers A, B, and C?



13.2. Are you confident in your answer at 13.1?
A. Sure
B. Not Sure

13.3. What is your reason for answering question 13.1?
A. Buoyant force is influenced by the portion of the object's volume that is submerged in the liquid
B. Buoyant force is influenced by the portion of the object's volume above the surface of the liquid
C. Buoyant force is influenced by the weight of the object
D. Buoyant force is influenced by the density of the object
E. ....
....

13.4. Are you confident in your answer at 13.3?
A. Sure
B. Not Sure

Figure 3. Example of MScI (Q13)

To ensure the quality and appropriateness of the four-tier instrument, a rigorous expert validation process was undertaken with the involvement of specialists in science education and assessment. Six science education experts looked at the MScI and looked at a number of things, such as whether the grammar was correct, whether the content matched the scientific concepts it was meant to, whether the material made sense, whether the reasoning in the third tier matched the answer choices in the first tier, and whether it fit with the conceptual frameworks that had been found. Table 4 shows the validation results that the six experts looked at. The examination of the validity test data reveals an average Content Validity Ratio (CVR) score of 0.966, surpassing the minimum acceptable threshold. This shows that the MScI is valid, so it can be used to assess pre-service elementary teachers' conceptual understanding of static fluid.

Table 4. Results of Recapitulation of MScI Validity by Experts

Aspects Validated	**Average Validity Score**
Readability of the questions	0,949
Suitability of the content to the concept	0,949
Logical coherence of the presented content	0,923
Congruence of third-tier reasoning with first-tier answer choice	1
Appropriateness in relation to the identified misconceptions	0,974
Average	0,966

The reliability of the MScI was assessed using the total scores obtained by each student. The reliability coefficient was determined to be 0.827 using trial data from 31 pre-service elementary school teachers. This value indicates a very high degree of reliability, per Guilford (1956). Additionally, the limited trial showed that participants understood the instrument items and that they produced consistent response patterns. It can therefore be concluded that the MScI is a stable and reliable diagnostic instrument.

Disseminate

The MScI was given to pre-service elementary teachers in the last step to see how well they understood static fluids. The distribution of the instrument facilitated the categorisation of each student's conception level based on the criteria delineated in Table 2. The resulting data are presented as percentage profiles of students' conception levels for static fluid concepts, as detailed in Table 5.

Table 5. Percentage of Students' Conception Level

Question Number	**Percentage of Students' Conception Level**				
					
Q1	25,8	25,8	32,3	16,1	0
Q2	16,1	3,2	45,2	25,8	9,7
Q3	12,9	12,9	54,8	12,9	6,5
Q4	9,7	25,8	51,6	12,9	0
Q5	6,5	25,8	54,8	12,9	0
Q6	0	12,9	58,1	29	0
Q7	0	32,3	51,6	16,1	0
Q8	0	22,6	51,6	25,8	0

Question Number	Percentage of Students' Conception Level				
					
Q9	22,6	32,3	35,5	9,7	0
Q10	0	9,7	48,4	41,9	0
Q11	0	19,4	58,1	22,6	0
Q12	6,5	35,5	41,9	16,1	0
Q13	6,5	22,6	58,1	9,7	3,2
Average	8,2	21,6	49,4	19,3	1,5

Table 5 analysis reveals significant differences in pre-service elementary teachers' conceptions. While some demonstrated a comprehensive understanding, others displayed only partial understanding, and some held misconceptions. A number of participants showed no understanding of static fluid concepts. These results show that the MScI can tell the difference between different levels of conceptual understanding in this area. Table 5 further reveals that, for several questions (Q6, Q7, Q8, Q10 and Q11), no participants achieved a sound understanding. By contrast, Question Q1 yielded the highest proportion of students classified at the sound understanding level relative to the other items. This finding indicates that numerous students possessed a solid understanding of hydrostatic pressure, particularly the principle that, in interconnected vessels, hydrostatic pressure is contingent upon depth rather than the height of the liquid above. Figure 4 gives more information about how common misconceptions are.

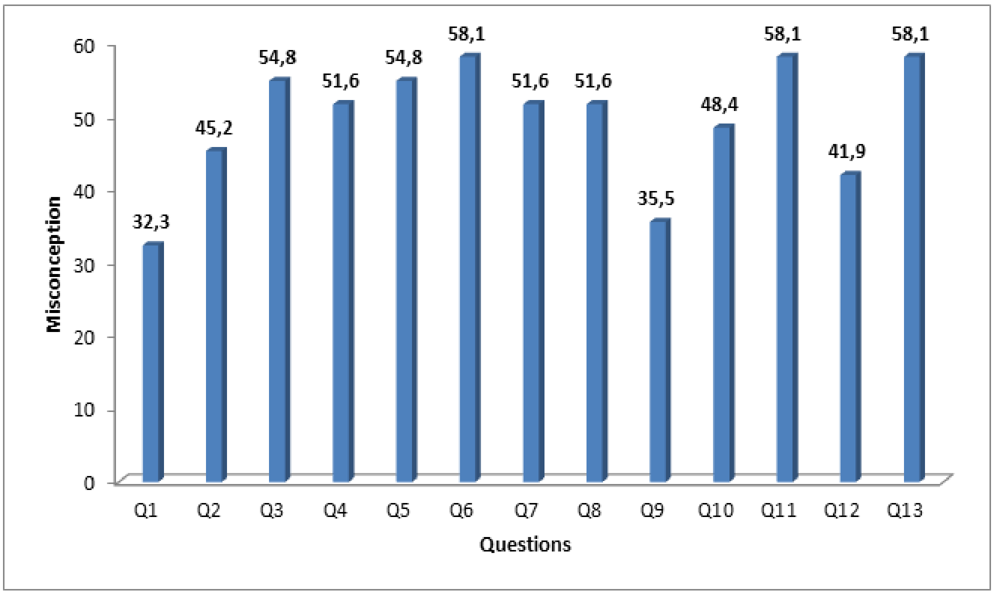


Figure 4. The percentages of pre-service elementary teachers' misconceptions

Figure 4 shows that pre-service elementary teachers had wrong ideas about every question in the MScI. The highest proportions of misconceptions were observed in Q6 and Q13, both of which concerned buoyant force, and in Q11, which addressed the position of objects in a liquid. In Question Q6, students assumed that the magnitude of the buoyant force is determined by the volume of liquid within the container. Students thought that the percentage of an object's volume above the liquid's surface affected the buoyant force acting on it in Question Q13. In fact, the magnitude of the buoyant force is determined solely by the density of the fluid, the volume of the object submerged within it, and the gravitational acceleration. Many pre-service teachers mistakenly believed that an object submerged horizontally would float in Question Q11, while an object submerged vertically or with a pointed end would sink. Figure 5 provides the context for Question Q11.

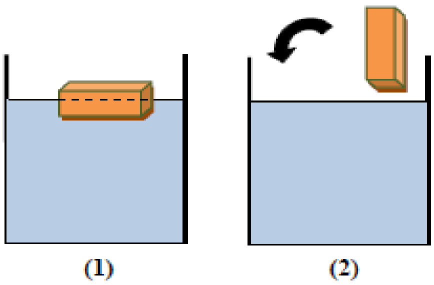


Figure 5. Context of the Illustration in Question 11

A block is horizontally submerged in a liquid and is seen to float in the context of Question Q11. Students were then asked to predict the outcome if the same block were submerged vertically in the same liquid. Many believed that the block would sink in the vertical position. Table 6 lists the specifics of the misconceptions found by the MScI.

Table 6. Students' misconceptions on the static fluids

Sub material	Question	Students' misconcepcions
Hydrostatic pressure	Q1	In a connected vessel, the amount of hydrostatic pressure depends on the height of the liquid above it.
	Q2, Q3	Hydrostatic pressure is affected by the volume of the space where the liquid is located
	Q4	The amount of hydrostatic pressure is not affected by the type of liquid
Buoyant force	Q5	The weight of an object in a liquid is less than the weight of an object in air
	Q6	The amount of buoyant force depends on the volume of liquid in the container
	Q7	The amount of buoyant force is affected by the viscosity of a liquid
	Q8	The amount of buoyant force on an object completely immersed in liquid is affected by the depth of the object in the liquid.
	Q13	The buoyant force on a floating object depends on the part of the object's volume above the surface of the liquid.
Position of objects in liquid	Q9	The heavier an object is, the more it will sink and the lighter an object is, the more it will float.
	Q10	A hollow solid object (reduced mass and volume) will change its position when immersed in a liquid.
	Q11, Q12	Objects dipped in a horizontal way will float, while objects dipped in avertical way or have sharp edges will sink.

DISCUSSIONS

The Define stage in this study aimed to create a solid theoretical and empirical basis for the development of the MScI. A thorough examination of the literature regarding various diagnostic test formats confirmed previous findings that four-tier diagnostic instruments yield superior diagnostic accuracy compared to two-tier or three-tier formats (Fratiwi et al., 2018; Kaltakci-Gurel et al., 2017; Kamilah et al., 2025). The decision to concentrate on static fluid concepts was influenced by their intrinsic conceptual complexity and the significant prevalence of related misconceptions documented in prior research (Surtiana et al., 2021; Xavier & Capellini, 2025; Zukhruf, 2018). Studies have consistently demonstrated that concepts such as hydrostatic pressure, buoyant force, and the positioning of objects in fluids are notably prone to misconceptions among pre-service teachers (Besson, 2004; Chattopadhyay, 2016; Goszewski et al., 2013 . These findings are consistent with the initial needs analysis conducted in the present study, which identified substantial gaps in understanding of these concepts among prospective elementary school teachers.

Employing a two-tier diagnostic instrument in the initial analysis was a crucial step in delineating students' conceptual frameworks. This finding is consistent with the research conducted by Kamcharean and Wattanakasiwich (2016) and Samsudin et al. (2017), which illustrated that the two-tier format effectively identifies reasoning patterns that can be further enhanced and validated in the development of the four-tier format. The reasoning data obtained at this stage provided an empirical foundation for constructing the third tier of the MScI, thereby ensuring that the distractors employed represent authentic misconceptions rather than hypothetical errors or assumptions derived from textbooks.

The Design Stage involved constructing the MScI using multiple forms of representation, including images, verbal explanations, diagrams, symbols, and graphs, to assess conceptual understanding more comprehensively. This strategy is in line with earlier research showing that integrating multiple representations can improve misconception detection accuracy (Aminuddin et al., 2024; Pramonoadi et al., 2020). The instrument was created using a four-tier structure. There were five possible answers in the first tier, including four distractions and one right response. The second tier measured confidence in the selected answer. The third tier contained semi-open reasoning options derived from the misconception analysis conducted during the Define stage, and the fourth tier measured confidence in the stated reasoning. Adding an open-ended reasoning option to the third tier (option E) made it possible to find unexpected student reasoning, which made the instrument's construct validity stronger. This design makes sure that the MScI tests not only the final answers but also how students think and how sure they are of their answers, both of which are important for finding out where students are wrong about static fluids.

The Develop stage led to the successful creation of thirteen MScI items, each consisting of the question stem, first-tier answer options, and third-tier reasoning options. Six experts in science education and assessment validated the instrument to make sure it was both high quality and relevant. The evaluation criteria encompassed linguistic accuracy, conceptual alignment, logical coherence of the material, consistency between the reasoning options in the third tier and the answer choices in the first tier, and alignment with the previously identified conception categories. The findings demonstrated the high validity (CVR = 0.966, exceeding the minimum threshold of 0.672) and high reliability (r = 0.827) of the Multitier Science Instrument (MScI) for identifying misconceptions in static fluids. These findings align with those documented for analogous diagnostic instruments in prior research (Kafiyani et al., 2019; Purwanto et al., 2020), thereby substantiating the robustness of the MScI in similar educational settings.

The high reliability of the MScI suggests that its multi-representational and multi-tier design effectively captures stable patterns in pre-service teachers' conceptual understanding. This result is consistent with evidence from systematic reviews showing that multi-tier diagnostic tools provide higher diagnostic precision than single-tier or two-tier alternatives, especially four-tier formats that integrate content, reasoning, and confidence (Ma et al., 2025). However, there is still little use of these tools in teacher education and elementary schools (Auliya et al., 2025).

The Disseminate stage evealed significant disparities in the conceptual comprehension of pre-service primary school teachers concerning static fluids, varying from a sound understanding to a total absence of comprehension. These results validate the MScI's ability to delineate levels of understanding comprehensively, aligning with the conclusions of Kaltakci-Gurel et al. (2017) and Samsudin et al. (2024), which indicated that multi-tier diagnostic tests surpass traditional instruments in detecting misconceptions. The absence of respondents in the sound understanding category for several items (Q6, Q7, Q8, Q10, Q11) highlights a fundamental weakness in the mastery of specific concepts, particularly buoyant force and the position of objects in fluids.

Many pre-service teachers still have false beliefs about buoyant force, linking it to either the volume of liquid in the container or the portion of an object above the liquid surface, as evidenced by the high percentage of misconceptions in Q6 and Q13. This finding aligns with prior research conducted by Kaltakci-Gurel et al. (2017) and Turgut and Gurbuz (2012), which indicated that such misconceptions frequently stem from quotidian intuitive experiences that contradict established physical laws. Misconceptions in Q11 about the position of objects in a fluid are similar to the common belief that an object's orientation affects its buoyancy, whereas the actual determining factors are its relative density and the volume submerged (Kiray et al., 2015). Taken together, these parallels confirm that deep-seated conceptual misunderstandings remain a persistent and widespread challenge.

CONCLUSION

This study reports on the development of the Multitier Science Instrument (MScI), a four-tier, multi-representational diagnostic tool created to identify and analyse misconceptions in static fluid concepts among pre-service elementary teachers. Its main innovation is how each test item incorporates verbal, visual, graphical, and symbolic forms of representation. This format makes it possible to conduct a thorough evaluation that documents students' final responses, the logic behind them, and their degree of confidence. The instrument achieved very high validity (CVR = 0.966) and excellent reliability (0.827), confirming its effectiveness as an assessment tool when compared with conventional single-tier tests. The findings showed that misconceptions occurred across all test items, with the highest prevalence in topics concerning buoyant force and the position of objects in liquids. These findings highlight the urgent need for science education to adopt more accurate diagnostic techniques. Despite its strengths, this study has limitations, most notably the modest sample size and the narrow focus on static fluids. As a result, the generalisability of the findings necessitates caution and should be investigated through subsequent studies involving larger and more diverse participant cohorts. Subsequent research may broaden the creation of analogous multi-representational, multitier diagnostic tools to additional physics concepts, thereby augmenting their utility for evaluating conceptual comprehension. Moreover, evaluating the MScI across varied educational contexts, including secondary education, may yield further insights and strengthen its broader relevance and utility. The findings of this study carry important implications for lecturers, teachers and curriculum developers. The MScI provides a robust basis for designing instructional strategies that are both targeted and pedagogically effective. The conceptual profiles produced by its application facilitate the creation of learning interventions that are precisely aligned with learners' actual levels of comprehension. Incorporating multi-representational, multitier assessments into teacher education programs has significant potential to enhance conceptual understanding, rectify enduring misconceptions, and bolster scientific literacy from the initial stages of education. This research significantly enhances the development of accurate and effective evaluation methodologies in science education, particularly in the context of training future elementary school teachers.

REFERENCES

Afif, N. F., Nugraha, M. G., & Samsudin, A. (2017, May). Developing energy and momentum conceptual survey (EMCS) with four-tier diagnostic test items. In *AIP Conference Proceedings* (Vol. 1848, No. 1, p. 050010). AIP Publishing LLC. https://doi.org/10.1063/1.4983966

Al Sultan, A., Henson, H., & Fadde, P. J. (2018). Pre-Service Elementary Teachers' Scientific Literacy and Self-Efficacy in Teaching Science. *IAFOR Journal of Education*, *6*(1). https://doi.org/10.22492/ije.6.1.02

Aminuddin, M., Salman, Z., Asim, & Irawati, A. (2024). Multi-representation approach in improving 1-dimensional kinematics conceptual understanding. *Universal Education Journal of Teaching and Learning*, *1*(2), 41–45. https://doi.org/10.63081/uejtl.v1i2.34

Auliya, H. R., Christianti, M., & Hidayat, B. (2025). Development of multi-tier diagnostic tests for primary schools: A systematic literature review. *Jurnal Prima Edukasia*, *13*(1), 59–74. https://doi.org/10.21831/jpe.v13i1.76906

Aydeniz, M., Bilican, K., & Kirbulut, Z. D. (2017). Exploring pre-service elementary science teachers' conceptual understanding of particulate nature of matter through three-tier diagnostic test. *International Journal of Education in Mathematics, Science and Technology*, *5*(3), 221–221. https://ijemst.com/index.php/ijemst/article/view/384

Besson, U. (2004). Students' conceptions of fluids. *International Journal of Science Education*, *26*(14), 1683–1714. https://doi.org/10.1080/0950069042000243745

Chattopadhyay, D. (2016). Classroom. *Resonance*, *21*(4), 381–386. https://doi.org/10.1007/s12045-016-0337-1

Kamilah, D. S., Aviyanti, L., Samsudin, A., & Suwarna, I. P. (2025, June). Pengembangan Digi4t-HB: Four-tier test digital untuk identifikasi miskonsepsi pada tekanan hidrostatis dan gaya apung. *In PROSIDING SEMINAR NASIONAL KEGURUAN DAN PENDIDIKAN* (SNKP) (Vol. 3, pp. 92-102). https://ejournal.ummuba.ac.id/index.php/SNKP/article/view/3116

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, *25*(6), 671–688. https://doi.org/10.1080/09500690305016

Eshach, H., Lin, T., & Tsai, C. (2018). Misconception of sound and conceptual change: A cross-sectional study on students' materialistic thinking of sound. *Journal of Research in Science Teaching*, *55*(5), 664–684. https://doi.org/10.1002/tea.21435

Fratiwi, N. J., Kaniawati, I., Suhendi, E., Suyana, I., & Samsudin, A. (2017, May). The transformation of two-tier test into four-tier test on Newton's laws concepts. In *AIP Conference Proceedings* (Vol. 1848, No. 1, p. 050011). AIP Publishing LLC. https://doi.org/10.1063/1.4983967

Fratiwi, N. J., Samsudin, A., & Costu, B. (2018). Enhancing K-10 Students' Conceptions through Computer Simulations-Aided PDEODE*E (CS-PDEODE*E) on Newton's Laws. *Jurnal Pendidikan IPA Indonesia*, *7*(2), 214–223. https://doi.org/10.15294/jpii.v7i2.14229

Galanis, C., Leske, M., Hamamura, T., Weber, N., Hing, N., Delfabbro, P. H., & King, D. L. (2025). Stigma in substance-based and behavioural addictions: A systematic review. *Journal of Behavioral Addictions*, *14*(1), 79-99. https://doi.org/10.1556/2006.2024.00072

Goszewski, M., Moyer, A., Bazan, Z., & Wagner, D. J. (2013, January). Exploring student difficulties with pressure in a fluid. In *Aip conference proceedings* (Vol. 1513, No. 1, pp. 154-157). American Institute of Physics. https://doi.org/10.1063/1.4789675

Guerra-Reyes, F., Guerra-Dávila, E., Naranjo-Toro, M., Basantes-Andrade, A., & Guevara-Betancourt, S. (2024). Misconceptions in the learning of natural sciences: A systematic review. *Education Sciences*, *14*(5), 497. https://doi.org/10.3390/educsci14050497

Guilford, J. P. (1956). *Fundamental Statistics in Psychology and Education* (3rd ed., Vol. 3191). McGraw-Hill.

Gurcay, D., & Gulbas, E. (2015). Development of three-tier heat, temperature and internal energy diagnostic test. *Research in Science & Technological Education*, *33*(2), 197–217. https://doi.org/10.1080/02635143.2015.1018154

Hermita, N., Suhandi, A., Syaodih, E., Samsudin, A., Isjoni, Johan, H., Rosa, F., Setyaningsih, R., Sapriadil, & Safitri, D. (2017). Constructing and Implementing a Four Tier Test about Static Electricity to Diagnose Pre-service Elementary School Teacher' Misconceptions. *Journal of Physics: Conference Series*, *895*, 012167. https://doi.org/10.1088/1742-6596/895/1/012167

Ilyas, A., & Saeed, M. (2018). Exploring teachers' understanding about misconceptions of secondary grade chemistry students. *International Journal for Cross-Disciplinary Subjects in Education*, *9*(1), 3323–3328. https://doi.org/10.20533/ijcdse.2042.6364.2018.0444

Kafiyani, F., Samsudin, A., & Saepuzaman, D. (2019). Development of a four-tier diagnostic test (FTDT) to identify students' mental models on static fluid. *Journal of Physics: Conference Series*, *1280*(5), 052030. https://doi.org/10.1088/1742-6596/1280/5/052030

Kaltakçi, D., & Didiş, N. (2007). Identification of Pre-Service Physics Teachers' Misconceptions on Gravity Concept: A Study with a 3-Tier Misconception Test. *AIP Conference Proceedings*, 499–500. https://doi.org/10.1063/1.2733255

Kaltakci-Gurel, D., Eryilmaz, A., & McDermott, L. C. (2017). Development and application of a four-tier test to assess pre-service physics teachers' misconceptions about geometrical optics. *Research in Science & Technological Education*, *35*(2), 238–260. https://doi.org/10.1080/02635143.2017.1310094

Kamcharean, C., & Wattanakasiwich, P. (2016). Development and Implication of a Two-tier Thermodynamic Diagnostic Test to Survey Students' Understanding in Thermal Physics. *International Journal of Innovation in Science and Mathematics Education*, *24*(2), 14–36. https://cmudc.library.cmu.ac.th/frontend/Info/item/dc:78702

Kiray, S. A., Aktan, F., Kaynar, H., Kilinc, S., & Gorkemli, T. (2015). A descriptive study of pre-service science teachers' misconceptions about sinking–floating. *Asia-Pacific Forum on Science Learning and Teaching*, *16*(2). https://eric.ed.gov/?id=EJ1096070

Laliyo, L. A. R., Botutihe, D. N., & Panigoro, C. (2019). The Development of Two-Tier Instrument Based On Distractor to Assess Conceptual Understanding Level and Student Misconceptions in Explaining Redox Reactions. *International Journal of Learning, Teaching and Educational Research*, *18*(9), 216–237. https://doi.org/10.26803/ijlter.18.9.12

Lombardi, D., Shipley, T. F., Bailey, J. M., Bretones, P. S., Prather, E. E., Ballen, C. J., Knight, J. K., Smith, M. K., Stowe, R. L., Cooper, M. M., Prince, M., Atit, K., Uttal, D. H., LaDue, N. D., McNeal, P. M., Ryker, K., St. John, K., van der Hoeven Kraft, K. J., & Docktor, J. L. (2021). The curious construct of active learning. *Psychological Science in the Public Interest*, *22*(1), 8–43. https://doi.org/10.1177/1529100620973974

Ma, H., Yang, H., Li, C., Ma, S., & Li, G. (2025). The effectiveness and sustainability of tier diagnostic technologies for misconception detection in science education: A systematic review. *Sustainability*, *17*(7), 3145. https://doi.org/10.3390/su17073145

Maryati, I., & Priatna, N. (2018). Analysis of statistical misconception in terms of statistical reasoning. *Journal of Physics: Conference Series*, *1013*, 012206. https://doi.org/10.1088/1742-6596/1013/1/012206

Menia, M., Mudzakir, A., & Rochintaniawati, D. (2017, May). The effect of conceptual metaphors through guided inquiry on student's conceptual change. In *AIP Conference Proceedings* (Vol. 1848, No. 1, p. 060009). AIP Publishing LLC. https://doi.org/10.1063/1.4983977

Mufida, S. N., Kaniawati, I., Samsudin, A., Suhendi, E., Aminudin, A. H., Umar, F. A., Astuti, I. R. W., Kunaedi, J., & Dewi, F. H. (2024). Developing Multitier Open-ended Transverse Wave Instrument (MOTWI): How to Assess Students' Misconceptions? *KnE Social Sciences*. https://doi.org/10.18502/kss.v9i13.16006

Ozkan, G., & Selcuk, G. S. (2015). The effectiveness of conceptual change texts and context-based learning on students' conceptual achievement. *Journal of Baltic Science Education*, *14*(6), 753–763. https://doi.org/10.33225/jbse/15.14.753

Ozkan, G., & Selçuk, G. S. (2012). How effective is "conceptual change approach" in teaching physics. *Journal of educational and Instructional Studies in the World*, *2*(2), 182-190. https://www.ajindex.com/dosyalar/makale/acarindex-1423906111.pdf

Pramonoadi, Tresnaningsih, S., Faqih, A., Setiani, R., & Dwikoranto. (2020). The validity of concept attainment with multi representation as an alternative learning model to improve students' mastery of concepts and scientific consistency. *Studies in Learning and Teaching*, *1*(2), 122–132. https://doi.org/10.46627/silet.v1i2.35

Purwanto, M. G., Suhandi, A., Coştu, B., Samsudin, A., & Nursanto, M. (2020). Static Fluid Concept Inventory (SFCI): A Gender Gap Analysis using Rasch Model to Promote a Diagnostic Test Instrument on Students' Conception. *International Journal of Advanced Science and Technology*, *29*(6), 3798–3812. https://avesis.yildiz.edu.tr/yayin/f11edcfe-88dd-4f0b-aa31-717aa7979d59/static-fluid-concept-inventory-sfci-a-gender-gap-analysis-using-rasch-model-to-promote-a-diagnostic-test-instrument-on-students-conception

Safitri, F., Rusdiana, D., Samsudin, A., & Widiyatmoko, A. (2025). Development of understanding test instruments for grade 7 junior high school students on temperature, heat and expansion topics. *Jurnal Penelitian Pendidikan IPA*, *11*(1), 395–404. https://doi.org/10.29303/jppipa.v11i1.9271

Samsudin, A., Suhandi, A., Rusdiana, D., Kaniawati, I., & Coştu, B. (2016). Investigating the effectiveness of an active learning based-interactive conceptual instruction (ALBICI) on electric field concept. *Asia-Pacific Forum on Science Learning and Teaching*, *17*(1). https://www.eduhk.hk/apfslt/v17_issue1/samsudin/page2.htm

Samsudin, A., Suhandi, A., Rusdiana, D., Kaniawati, I., & Coştu, B. (2017). Promoting Conceptual Understanding on Magnetic Field Concept Through Interactive Conceptual Instruction (ICI) with PDEODE*E Tasks. *Advanced Science Letters*, *23*(2), 1205–1209. https://doi.org/10.1166/asl.2017.7539

Samsudin, A., Zulfikar, A., Saepuzaman, D., Suhandi, A., Aminudin, A. H., Supriyadi, S., & Coştu, B. (2024). Correcting grade 11 students' misconceptions of the concept of force through the conceptual change model (CCM) with PDEODE*E tasks. *Journal of Turkish Science Education*, *21*(2), 212–231. https://doi.org/10.36681/tused.2024.012

Silva Mangiante, E. (2025). Addressing Misconceptions About Plant Growth. *Science and Children*, *62*(1), 73–79. https://doi.org/10.1080/00368148.2024.2437194

Suhandi, A., Surtiana, Y., Husnah, I., Setiawan, W., Siahaan, P., Samsudin, A., & Costu, B. (2020). Fostering high school students' misconception about boiling concept using conceptual change laboratory (CCLab) activity. *Universal Journal of Educational Research*, *8*(6), 2211–2217. https://doi.org/10.13189/ujer.2020.080603

Suprapto, N. (2020). Do we experience misconceptions?: An ontological review of misconceptions in science. *Studies in Philosophy of Science and Education*, *1*(2), 50–55. https://doi.org/10.46627/sipose.v1i2.24

Surtiana, Y., Suhandi, A., Putri, K., Setiawan, W., Siahaan, P., & Samsudin, A. (2021). Penerapan model virtual conceptual change laboratory (VCCLAB) untuk meremediasi miskonsepsi siswa pada materi tekanan hidrostatik [The implementation of the Virtual Conceptual Change Laboratory (VCCLAB) model to remediate students' misconceptions on hydrostatic pressure material]. *WaPFi (Wahana Pendidikan Fisika)*, *6*(1), 63–70. https://doi.org/10.17509/wapfi.v6i1.32451

Turgut, U., & Gurbuz, F. (2012). Effect of conceptual change text approach on removal of students' misconceptions about heat and temperature. *International Journal of Innovation and Learning*, *11*(4), 386. https://doi.org/10.1504/IJIL.2012.047139

Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the Critical Values for Lawshe's Content Validity Ratio. *Measurement and Evaluation in Counseling and Development*, *45*(3), 197–210. https://doi.org/10.1177/0748175612440286

Xavier, I. N., & Capellini, S. A. (2025). Remediation program with working memory and reading for students with learning difficulties: elaboration and pilot study. *Children*, *12*(4), 426. https://doi.org/10.3390/children12040426

Yalçin, S. A., Yalçin, P., Akar, M. S., & Sağirli, M. Ö. (2017). The effect of teaching practices with real life content in light and sound learning areas. *Universal Journal of Educational Research*, *5*(9), 1621–1631. https://doi.org/10.13189/ujer.2017.050920

Yan, Y. K., & Subramaniam, R. (2018). Using a multi-tier diagnostic test to explore the nature of students' alternative conceptions on reaction kinetics. *Chemistry Education Research and Practice*, *19*(1), 213–226. https://doi.org/10.1039/C7RP00143F

Yeo, J.-H., Yang, H.-H., & Cho, I.-H. (2022). Using a three-tier multiple-choice diagnostic instrument toward alternative conceptions among lower-secondary school students in taiwan: taking ecosystems unit as an example. *Journal of Baltic Science Education*, *21*(1), 69–83. https://doi.org/10.33225/jbse/22.21.69

Yulianawati, D., Muslim, Hasanah, L., & Samsudin, A. (2018). A case study of analyzing 11 th graders' problem solving ability on heat and temperature topic. *Journal of Physics: Conference Series*, *1013*, 012042. https://doi.org/10.1088/1742-6596/1013/1/012042

Zukhruf, K. (2018). Mengidentifikasi miskonsepsi fluida statis pada mahasiswa calon guru fisika universitas Samudra [Identifying misconceptions of static fluid among pre-service physics teachers at Samudra University]. *Gravitasi*, *1*(02), 11–16. https://media.neliti.com/media/publications/394982-none-958aff65.pdf