ORBITA: Jurnal Pendidikan dan Ilmu Fisika
ISSN : 2460-9587 (print) | ISSN : 2614-7017 (elektronik)
Volume 11, No. 1, May 2025, pp. 01-11

Design and validation of a diagnostic test instrument for identifying misconceptions in basic physics

Zahriah[1*], Fera Annisa[1], Nur Hidayah Binti Ahmad[2], Adinda Fahira[1]

[1]Department of Physics Education, Universitas Islam Negeri Ar-Raniry Banda Aceh, Aceh, Indonesia
[2]Doctor of Physics, Universiti Teknologi Malaysia, Johor, Malaysia
*Correspondence: zahriah.ilyas@ar-raniry.ac.id

Received: 08 November 2024 | Revised: 17 February 2025 | Accepted: 17 February 2025 | Published Online: 10 May 2025

Abstract

Misconceptions are a classic problem experienced by students in learning which can affect academic achievement, and difficulty understanding advanced physics concepts, and this has the potential to inherit the same mistakes when students become teachers. The unavailability of instruments to identify misconceptions makes it difficult for lecturers to explore student misconceptions, so it is necessary to develop a set of diagnostic test instruments. This research aims to design and test the feasibility of a three-level diagnostic test instrument which will then be used as a tool to identify student misconceptions in the Basic Physics course. This research method is research and development (R&D) of the Plomp model with stages: 1) preliminary investigation, 2) design, 3) realization/construction, 4) testing, evaluation, and revision, and 5) implementation. Data collection techniques use questionnaires and tests, with research instruments in the form of validation sheets and diagnostic test questions. The content validity data analysis technique uses a percentage scale of feasibility categories based on expert judgment and construct validity testing using item analysis software version 2.03. The results of testing the feasibility of the diagnostic test instrument in terms of content validity obtained an overall percentage of 97.21%, while the construct validity test found that 75% of the questions were suitable for use. Thus, it can be concluded that the diagnostic test instrument that has been designed is suitable for use and makes a significant contribution to identifying student misconceptions, as well as helping to improve the quality of education, especially in physics learning, and supporting better academic goals.

Keywords: diagnostic test; misconceptions; basic physics

How to Cite: Zahriah, Annisa, F., Ahmad, N. H. B., Fahira, A. (2025). Design and validation of a diagnostic test instrument for identifying misconceptions in basic physics. *ORBITA: Jurnal Pendidikan dan Ilmu Fisika*, *11*(1), 01-11. https://doi.org/10.31764/orbita.v11i1.27782

INTRODUCTION

Misconceptions are a universal problem that occurs in learning, including in science learning and this is a classic problem that requires serious attention, especially at the university level. Misconceptions are often not detected properly by students themselves or by lecturers as teachers, so this problem can continue throughout the learning process (Izza et al., 2021). If in studying the most basic physics material students have experienced misconceptions, then the next material will also bring wrong concepts (Raflah et al., 2021).

There are several reasons for misconceptions among students, including (1) misunderstandings that were previously formed from previous student learning experiences (Ananda & Syuhendri, 2021), (2) incomplete or partial understanding of a concept, (3) errors in interpreting and understanding a concept, 4) poor reading skills, 5) unclear content of teaching materials, and 5) misconceptions caused by regional language where sometimes terms in regional language are different from the meaning of scientific language. Apart from that, misconceptions are also caused by misinterpreting physical phenomena that occur in everyday life which are not in line with scientific knowledge (Neidorf et al., 2019), influences from within students, educators, the language used, characteristics of teaching materials, and reference book (Resbiantoro et al., 2022). Suppose lecturers or instructors are aware of and understand the causes of misconceptions among students. In that case, they can take appropriate steps to improve the learning process and quality so that they get the expected results.

Misconceptions are one of the causes of a lack of academic achievement and problem-solving skills in many fields of science, technology, engineering, and mathematics (Achor et al., 2022). This statement is supported by data on student learning outcomes from the Physics Education Study Program at UIN Ar-Raniry Banda Aceh, where it was found that there were still many students who got grades below the class average, namely 64.70% in 2022 and 50% in 2023. This indicates that students lack strong basic concepts and allows students to experience conceptual errors in learning the basics of physics.

This misconception not only impacts basic understanding but also affects students' ability to apply concepts critically and creatively in solving more complex problems (Wiyartiningtyas & Haryani, 2023). For example, conceptual errors in basic physics can hinder understanding of more advanced material, such as thermodynamics or quantum mechanics. Without a correct and in-depth knowledge of concepts, students will have difficulty connecting the concepts learned and applying them to real phenomena, thereby reducing the effectiveness of overall learning outcomes. Therefore, lecturers or instructors need to identify these misconceptions through appropriate steps, one of which is by designing diagnostic test instruments. Diagnostic tests are used to determine students' weaknesses in learning, especially in understanding basic concepts which are often a source of misconceptions (Sutiah, 2020). Through this test, lecturers can identify areas that are still poorly mastered or misunderstood by students, thus allowing for more precise and focused treatment. Diagnostic tests also provide a more detailed picture of student’s initial knowledge before starting further learning, so that they can help design learning strategies that suit student needs (Yusrizal & Rahmati, 2020).

Diagnostic test instruments have been developed for chemistry subjects at the high school level (Nisa & Sudrajat, 2023), mathematics subjects (Damayanti & Priatmoko, 2023), high school physics subjects (Wahyudi et al., 2021), and biology subjects (Tika et al., 2023). Diagnostic test instruments have also been developed for prospective elementary school teacher students (Yulianawati et al., 2022), and in fluid mechanics courses (Nurhafsari & Rismaningsih, 2023). Although diagnostic test instruments have been developed by previous researchers for various subjects and levels of education such as chemistry, mathematics, and biology at the secondary school level, as well as courses at the university level such as fluid mechanics, no researcher has yet developed a diagnostic test instrument specifically for Basic Physics. The Basic Physics course is one of the main foundations in science education, especially for physics education program students. They are prospective teachers who will teach physics subjects at school after they graduate. Of course, they must have correct and in-depth basic physics concepts, as well as support the learning of advanced physics material. The unavailability of appropriate diagnostic test instruments can hinder the effective identification of misconceptions so that the learning strategies implemented are not fully able to increase students' understanding of concepts optimally.

This research aims to design and test the validity of a three-tier diagnostic test instrument which will later be used to identify misconceptions experienced by student teachers in the Basic Physics course. The diagnostic test instrument is expected to be able to explore specific problems of conceptual errors that prospective physics teacher students have in basic materials in the Basic Physics course. This instrument is also designed to provide accurate and measurable data regarding students' conceptual understanding, thereby enabling lecturers or tutors to design more targeted learning. Thus, the use of this instrument is not only aimed at evaluating student understanding, but also as a tool in the process of improving basic physics learning to make it more effective and efficient, and can be used by lecturers or physics teachers in other study programs that require a basic understanding of physics as a foundation.

METHODS

This research uses the research and development (R&D) method with the Plomp model which consists of five stages (Plomp & Nieveen, 2010). The stages include: (1) Preliminary Investigation, where the initial analysis is conducted to identify problems, needs, and the theoretical basis for developing the diagnostic test instrument; (2) Design, which involves designing the diagnostic test instrument framework, creating a blueprint or grid, and drafting the initial version of the three-tier test; (3) Realization/Construction, where the test instrument is developed based on the design and reviewed for validity by experts; (4) Testing, Evaluation, and Revision, which includes testing the instrument on a limited group of students, analyzing the results, and revising the instrument to ensure its reliability and effectiveness; and (5) Implementation, where the final instrument is applied to a broader group of students to evaluate its practicality and impact in identifying misconceptions in the Basic Physics course. The following Figure 1 shows the steps of the Plomp model development method.

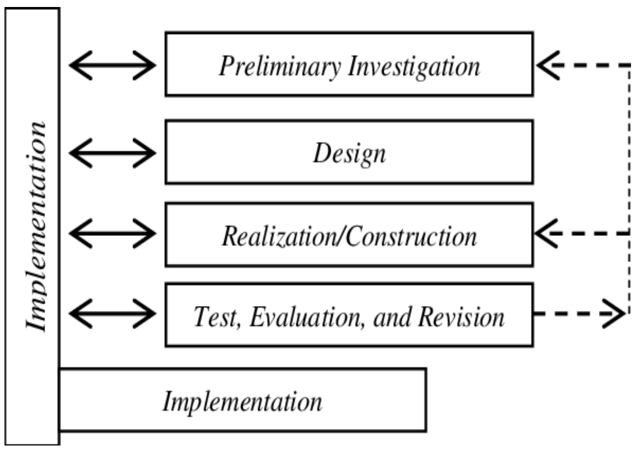


Figure 1. Plomp Model R&D Method

In this research, the implementation stage was not carried out due to time constraints and the scope of the study, which focused primarily on the development and validation of the diagnostic test instrument. The research concluded at the testing, evaluation, and revision stage, where the instrument's validity, reliability, and ability to identify misconceptions were rigorously analyzed. This ensures that the instrument is ready for future implementation in broader settings, allowing subsequent research to explore its practical application and effectiveness in real classroom environments.

The data collection techniques used were questionnaires and tests. Questionnaires are used to obtain data regarding the feasibility of diagnostic test instruments that have been developed based on the review and assessment of validators (expert judgment), while tests are used to obtain data regarding the appropriateness of question items that have been developed based on the results of small-scale trials. The sample selection in this research was carried out using a purposive sampling technique, namely sampling with certain considerations (Sugiyono, 2017), where the researcher selected validators based on their educational background and field of expertise, while the trial sample chosen was 24 students who were or had took a Basic Physics course. Before selecting the sample, the researcher first asked the validators and students to pay attention to research principles and ethics. The data analysis techniques used in this research are: 1) content validity analysis which is based on the results of the validation sheet assessment completed by the validators. To determine the percentage category for the instrument's feasibility level, use the following criteria in Table 1.

Table 1 Instrument Feasibility Percentage Category (Saputra, 2021)

Percentage	Category
<20%	Not Feasible
21% - 40 %	Not Worth It
41% - 60%	Decent Enough
61% - 80%	Worthy
81% - 100%	Very Worthy

2) Construct validity analysis: carried out to ensure that the questions used in the test can measure ability or knowledge accurately, validly, and reliably. The data obtained was processed using *item analysis software version 2.03* to determine the quality of the questions in terms of validity, reliability, difficulty, and distinguishing power.

RESULT AND DISCUSSIONS

Design of Diagnostic Test Instruments

The diagnostic test instrument designed in this research was carried out in several stages, namely: 1) Carrying out a curriculum analysis, namely reviewing the suitability of the study materials for the Basic Physics course with the learning outcomes of the course (CLO) and sub-CLO; 2) Develop a diagnostic test instrument grid by formulating question indicators that are relevant to the sub-CLO and study materials and determining the appropriate cognitive level. 3) create question items in the form of three-tier multiple choice. The first level is a multiple choice question consisting of 5 choice options (A, B, C, D, and E), the second level is the reason for choosing the answer at the first level, while the third level is the level of confidence in choosing the answer. The grid display of the test instrument and the display of one of the questions developed are presented in Figure 2 and Figure 3.

DIAGNOSTIC TEST INSTRUMENT GRID

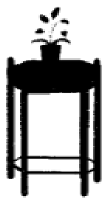
Objective : Identify misconceptions
Course : Basic Physics
Number of Questions : 40 Items

COURSE LEARNING OUTCOMES (CLO)	
1	Mastering basic physics concepts about matter, energy and structure of substances, as well as the application of physics in technology
2	Apply basic physics concepts and appropriate mathematical methods to obtain solutions to quantitative problems in physics.
3	Able to formulate a simple physical system into a physical model and complete the physical model with the help of mathematics
4	Able to work independently effectively and collaborate in lecture assignment groups

No	Sub-CLO	Study Materials	Indicators	Cognitive Level	Question Number	Question Form
1	Students can understand the concept of measurement and use measuring instruments correctly	Quantity and Measurement	Given illustrations of measurement activities, students are asked to determine the appropriate type of measuring instrument accompanied by appropriate reasons	C2	1	Multiple Choice Three-Tier Test
			Presented with pairs of physical quantities, students are asked to determine equivalent pairs of quantities accompanied by appropriate reasons	C2	2	Multiple Choice Three-Tier Test

Figure 2. Instrument Grid View

1. A flower pot is placed motionless on a table as shown in the following picture. What happens to the flower pot is...
   A. The gravitational force is greater than the normal force
   B. The friction force of the pot against the table surface is very large
   C. The external force acting is too small
   D. The magnitude of the gravitational force is the same as the normal force
   E. No forces are acting



Reasons for choosing an answer:
A. An immovable object is always influenced by a force, where the resultant force is zero
B. The object remains motionless because the external force is smaller than the force possessed by the object
C. The large friction force causes the object not to move
D. Objects do not move because the gravitational force is greater, thus preventing the object from moving
E. The object does not move because there is no force acting on the object

Answer confidence level:

1	2	3	4	5	6
Just guessing	Very unsure	Not sure	Sure	Very sure	Very, very sure

Figure 3. Display of one of the question items

Figure 3 above shows that the diagnostic test questions are arranged in three levels, which are intended to measure gradually and in-depth the student's level of understanding of an idea. In the first stage, students can only choose answers from the available options. In the second stage ask students to explain why they answered that way. At this level is very important to know whether students' understanding is truly deep or whether misconceptions need to be corrected. Next, the third level aims to assess students' level of confidence in the answers chosen at the first level. With this, teachers can see whether students are confident in their understanding or perhaps just guessing.

Feasibility Assessment by Expert Judgment

The design of the diagnostic test instrument that has been developed is very suitable for use as a tool for identifying misconceptions. This is shown by the assessment results of experts in the field of physics material and experts in the field of evaluation of physics education which are presented in the following table.

Table 2. Physics Material Expert Assessment Results

Aspect	V 1 (%)	V 2 (%)	V 3 (%)	Average	Eligibility Level
Continuity of question items with CPMK and Sub CPMK	90,00	89,35	91,75	90,36	Very worthy
Conformity of question items to basic physics concepts	93,63	90,25	90,46	91,45	Very worthy
Accuracy of concepts and answers	85,90	88,32	90,00	88,07	Very worthy
Depth of material and level of difficulty	90,35	90,00	91,25	90,53	Very worthy
Ability to identify misconceptions	90,50	91,00	92,53	91,34	Very worthy
Clarity of the language used	92,25	90,56	91,20	91,35	Very worthy
Clarity of question item construction	93,40	92,25	92,65	92,76	Very worthy
Average Percentage				**90,84**	Very worthy

The designed assessment instrument meets high suitability standards in the field of physical materials, as shown in Table 2 above. This instrument can be said to be suitable for use as a measurement or evaluation tool in physics learning, with an overall average percentage of 90.84% in the very appropriate category. The high level of appropriateness in all aspects of the assessment shows that this instrument is not only relevant to the material being measured but is also designed appropriately and uses language that is easy to understand. As a result, it is hoped that this instrument can provide accurate and reliable results for evaluating students' or respondents' understanding of the physics material being tested. The instrument's success in meeting the very appropriate category also provides additional confidence that it can be used well in a variety of contexts, such as in teaching, research, and assessment. The following table shows the assessment of expert validators in the field of physics education evaluation.

Table 3. Physics Education Evaluation Expert Assessment Results

Aspect	Indicator	V 1 (%)	V 2 (%)	Average	Eligibility Level
Material/Content	• Suitability of question items to learning outcomes • Logical and homogeneous answer choices • Presentation of answer choices that are relevant to the question	97,04	96,96	97,00	Very worthy
Construction	• Clarity of instructions for working on	97,06	97,63	97,34	Very worthy

Aspect	Indicator	V 1 (%)	V 2 (%)	Average	Eligibility Level
	questions • Strictness in the formulation of the main questions • The question stem does not provide clues to the answer • The main question does not provide a double negative statement • The length of the answer choices is relatively the same • Clarity of the chronology of answer choices • Clarity of presentation of images/diagrams • The independence of the question item from the previous question				
Language	• Language compatibility with the KBBI • Use of communicative language • Use of clear sentences • Do not use regional languages	96,56	97,50	97,03	Very worthy
Average Percentage				**97,11**	**Very worthy**

The results of the two validators' assessments showed that the diagnostic test equipment met the eligibility standards in all aspects evaluated. This instrument can be said to be ready to be used effectively to measure diagnostic capabilities because it has an average percentage of 97.11% in the very feasible category. While a high score on the construction aspect shows that the questions are well designed and under the principles of correct instrument preparation, a high score on the material/content aspect shows that the question items are by the competency or material to be measured. Apart from that, the language elements assessed were also very good, showing that the instrument was easy to understand and did not leave participants ambiguous.

Diagnostic Test Instrument Trial

Small-scale trials were carried out to test the construct validity of each question. Next, the trial data was analyzed using item analysis software version 2.03 to evaluate the validity, reliability, level of difficulty, and distinguishing power. For each question, the following results were obtained:

Table 4 Results of Instrument Trials on a Small Scale

Validity		**Difficulty level**		**Differentiating power**		**Reliability (KR-20)**	
Category	**%**	**Category**	**%**	**Category**	**%**	**Value**	**Category**
Very high	0,00	Easy	7,50	Very good	0,00	0,734	High
High	15,00	Currently	47,50	Good	10,00		
Enough	70,00	Difficult	45,00	Enough	67,50		
Low	2,50			Bad	10,00		

Validity		**Difficulty level**		**Differentiating power**		**Reliability (KR-20)**	
Category	**%**	**Category**	**%**	**Category**	**%**	**Value**	**Category**
Very low	12,50			Throw away	12,50		

Decision	**Amount**	**Percentage**
Question items are valid	34	85 %
The question item is invalid	6	15%

For each assessment aspect analyzed using item analysis software version 2.03, the results were: 1) the percentage of test item validity was 0.00% very high, 15.00% high, 70.00% sufficient, 2.50% low, and 12. 50% is very low; 2) The percentage of different power of questions in the very good category is 0.00%, good 10%, fair 67.50%, bad 10.00%, and discarded 12.50%; 3) The difficulty level of questions in the easy category is 7.50%; 4) for the reliability of the questions using the KR-20 test, a value of 0.734 was obtained in the high category. Thus, overall of the 40 questions developed, 87.5% were declared valid and the remaining 12.5% were declared invalid, so 34 questions were suitable for use and the remaining 6 were not used. Therefore, these questions were removed to maintain the quality of the measurements, so that the results produced are more accurate in finding misconceptions.

Most of the questions are valid (85%), which shows that the quality of the question items is quite good based on validity, and there are still several items (15%) that need to be improved. For the level of difficulty, most of the questions were in the medium difficulty level category (47.50%), which is ideal for assessing participants' abilities fairly. The percentage of difficult questions is quite high (45.00%), indicating that the questions are challenging. However, the difficulty level of the questions has been set so that it is not too difficult so that it does not hinder the achievement of the assessment objectives. Questions that are classified as easy are only 7.50%. Therefore, the variety of easy questions needs to be increased to create a more balanced distribution of difficulty levels. In terms of discriminating power, most of the questions were in the sufficient category (67.50%). Only 10% of questions have good discriminating power, and there are no questions with very good discriminating power.

As many as 22.50% of the questions (consisting of the bad and discarded categories) need to be revised or deleted because they have low differentiating power. The KR-20 value shows that the internal consistency of the test is quite good. This test reliably measures the intended goals, although there is still room for improvement. The very high category has a validity value of 0.00%. This is likely due to limited variation in respondents, such as the ability of test participants which tends to be homogeneous, so that it does not allow for a very high correlation between items and the total score. Apart from that, the homogeneity of participants' abilities also causes the differentiating power of questions to tend to be low, even on questions that are actually of high quality (Mardapi, 2019).

The results of this study have a significant impact on the future development of diagnostic tests intended to uncover misconceptions. It is important to use valid, accurate, and useful tests to spot misconceptions in Basic Physics courses early on. This test not only helps lecturers or educators find misconceptions, but also provides a strong basis for creating better learning strategies, using better teaching materials, and choosing the right media to help students learn.

Previous studies show that using diagnostic tests as learning assessments improves student learning outcomes compared to a series of summative tests. This is because using diagnostic tests makes it possible to identify learners' weaknesses in understanding the material and then provide solutions to improve them (Esomonu & Eleje, 2020). Diagnostic tests are also very effective for identifying student needs. Educators can also find ways to improve their academic abilities and help students who experience learning difficulties (Conforme et al., 2019). By carrying out diagnostic assessments when implementing the independent curriculum in Indonesia, educators can identify students' knowledge and psychological gaps through initial assessments. This allows them to create learning activity programs that focus on the most important classroom issues (Nugroho et al., 2023). Furthermore, three levels of diagnostic tests can be used to find out about the misconceptions that occur, the types of misconceptions experienced by students, and the factors that cause these misconceptions to arise (Maryam, 2020).

The designed diagnostic test instrument was tested on a limited sample, thereby affecting the generalization of research findings. Although this diagnostic test instrument shows valid content and constructs, it is possible that some things are not well identified. As a result, the question items must be revised to find various misconceptions in the Basic Physics course more thoroughly. In addition, the test format used in this research does not fully utilize the latest technology which is more interactive and responsive, which can produce more accurate and efficient results.

Taking these limitations into account, the researchers suggest further research to apply the test to larger and more diverse samples. This will ensure that this diagnostic test can be widely used and takes into account external factors that may influence test results. In addition, further validation needs to be carried out using more diverse methods and involving more experts to ensure that the results are consistent. This is because misconceptions are not only a problem at one level or course. This is because misconceptions are not just a problem that occurs in one subject or level of education; they can occur at various levels of education and fields of study, so it is important to develop tools that can detect misconceptions in a variety of situations. Comprehensive diagnostic tests can help teachers, both lecturers and teachers at various levels, choose the right intervention strategies to improve understanding of concepts.

CONCLUSION

The diagnostic test instrument developed in this research was declared very feasible based on the assessment of physics material experts with a percentage of 90.84%, and was also declared very feasible based on the assessment of physics education evaluation experts with a percentage of 97.11%, with details of material/content aspects 97.00%, construction aspect 97.34%, and language aspect 97.03%. Meanwhile, based on the results of the construct validity test using item analysis software version 2.03, there were 6 questions (15%) that could not be used as diagnostic test instruments to identify misconceptions because they did not meet the eligibility requirements. Thus, the number of questions that can be used is 34 items (85%).

REFERENCES

Achor, E. E., Ellah, B. O., & Omaga, J. O. (2022). Misconceptions and Difficult Concepts as Determinant of Students’ Academic Engagement and Retention in Physics. *Jurnal VARIDIKA*, *1*(1), 42–52. https://doi.org/10.23917/varidika.v1i1.17660

Ananda, L., & Syuhendri, S. (2021). Miskonsepsi mahasiswa calon guru fisika pada mata kuliah ipba materi periode orbit bulan mengelilingi bumi. *Prosiding Seminar Nasional Pendidikan IPA Tahun 2021*.

Conforme, D. F. I., Romero, A. L. C., Romero, D. C., & Laz, E. M. S. A. (2019). Application of diagnostic assessment on beginning school year. *International Research Journal of Management, IT and Social Sciences*, *6*(5), 53–59. https://doi.org/10.21744/irjmis.v6n5.701

Damayanti, Y. D., & Priatmoko, S. (2023). Pengembangan instrumen tes diagnostik three-tier multiple choice test (ttmct) untuk menganalisis miskonsepsi siswa pada konsep sifat keperiodikan unsur. *Chemined*, *12*(2), 125–130. http://journal.unnes.ac.id/sju/index.php/chemined

Esomonu, N. P.-M., & Eleje, L. I. (2020). Effect of Diagnostic Testing on Students' Achievement in Secondary School Quantitative Economics. *World Journal of Education*, *10*(3), 178. https://doi.org/10.5430/wje.v10n3p178

Izza, R. I., Nurhamidah, N., & Elvinawati, E. (2021). Analisis Miskonsepsi Siswa Menggunakan Tes Diagnostik Esai Berbantuan Cri (Certainty of Response Index) Pada Pokok Bahasan Asam Basa. *Alotrop*, *5*(1), 55–63. https://doi.org/10.33369/atp.v5i1.16487

Mardapi, D. (2019). *Psikologi Pengukuran Pendidikan*. PT Remaja Rosdakarya.

Maryam, E. (2020). Identifikasi Miskonsepsi Menggunakan Three-Tier Diagnostic Test Berbasis Google Form pada Pokok Bahasan Potensial Listrik. *Silampari Jurnal Pendidikan Ilmu Fisika*, *2*(2), 149–162. https://doi.org/10.31540/sjpif.v2i2.1083

Neidorf, T., Arora, A., Erberber, E., Tsokodayi, Y., & Mai, T. (2019). Student Misconceptions and Errors in Physics and Mathematics. In *IEA Research for Education* (Vol. 9).

Nisa, K., & Sudrajat, A. (2023). Pengembangan Instrumen Tes Diagnostik Five-Tier untuk Mengidentifikasi Miskonsepsi Siswa Kelas XI pada Materi Laju Reaksi. *PENDIPA Journal of Science Education*, *7*(2), 127–136. https://doi.org/10.33369/pendipa.7.2.127-136

Nugroho, D., Wirawan, W., Febriantania, P., & Ridaningsih, I. (2023). A Sistematic Literature Review : Implementasi Asesmen Diagnostik pada Kurikulum Merdeka. *Jurnal Pendidikan Islam*, *9*(2), 50–61. https://doi.org/10.37286/ojs.v9i2.197

Nurhafsari, A., & Rismaningsih, F. (2023). Pengembangan Instrumen Four Tier Diagnosti c Test Berbasis iSpring Suite 9 Untuk Mengidentifikasi Pemahaman Konsep Pada Mahasiswa. *Jurnal Pendidikan Borneo (Borju)*, *5*(2), 245–259. https://garuda.kemdikbud.go.id/documents/detail/3692163

Plomp, T., & Nieveen, N. M. (2010). *An introduction to educational design research*.

Raflah, N., Salasi, & Umam, K. (2021). Analisis Miskonsepsi Siswa pada Materi Logaritma di SMA Negeri 5 Banda Aceh. *Jurnal Peluang*, *9*(2), 1–8. https://doi.org/10.24815/jp.v9i2.26339

Resbiantoro, G., Setiani, R., & Dwikoranto. (2022). A Review of Misconception in Physics: The Diagnosis, Causes, and Remediation. *Journal of Turkish Science Education*, *19*(2), 403–427. https://doi.org/10.36681/tused.2022.128

Saputra. (2021). *Pengembangan Bahan Ajar Sejarah Berbasis WEB*. YLGI.

Sugiyono. (2017). *Metode Penelitian Pendidikan Pendekatan Kuantitaif, Kualitatif, dan R&D*. Alfabeta.

Sutiah. (2020). *Optimalisasi Fuzzy Topsis: Kiat meningkatkan prestasi belajar mahasiswa*. Nizamia Learning Centre.

Tika, P. N., Jariah, Y. A., Melina, M. M., Ristanto, R. H., & Isfaeni, H. (2023). Pengembangan Instrumen Tes Diagnostik Three-Tier Pada Pembelajaran Sistem Ekskresi Berdiferensiasi. *Bio-Lectura : Jurnal Pendidikan Biologi*, *10*(2), 167–182. https://doi.org/10.31849/bl.v10i2.14610

Wahyudi, F., Didik, L. A., & Bahtiar, B. (2021). Pengembangan Instrumen Three Tier Test Diagnostik Untuk Menganalisis Tingkat Pemahaman Dan Miskonsepsi Siswa Materi Elastisitas. *Relativitas: Jurnal Riset Inovasi Pembelajaran Fisika*, *4*(2), 48. https://doi.org/10.29103/relativitas.v4i2.5184

Wiyartiningtyas, L. P., & Haryani, F. F. (2023). *The Advantages of Diagnostic Tests for High School Students in Physics Learning: A Literature Review*. Atlantis Press SARL.

https://doi.org/10.2991/978-2-38476-060-2_10

Yulianawati, D., Wahyuningsih, A., & Pebriana, N. A. (2022). Pengembangan Instrumen 4TSDT (Four Tier – Science Diagnostic Test) untuk Mengidentifikasi Level Konsepsi Calon Guru Sekolah Dasar. *Jurnal Basicedu*, *6*(6), 9483–9490. https://doi.org/10.31004/basicedu.v6i6.4117

Yusrizal, & Rahmati. (2020). *Tes Hasil Belajar*. Bandar Publishing.