JPPIPA 11(11) (2025)

Jurnal Penelitian Pendidikan IPA

http://jppipa.unram.ac.id/index.php/jppipa/index

Empirical Validation of a Conceptual Understanding Test on Dynamic Electricity for High School Students

Karolus Boromeus Mura[1], Sujito[1*], Ahmad Taufiq[1], Ivan Danar Aditya Irawan[1], Nur Akhyar Basri[1]

[1] Department of Physics, State University of Malang, Malang, Indonesia.

Received: August 05, 2025
Revised: October 18, 2025
Accepted: November 25, 2025
Published: November 30, 2025

Corresponding Author:
Sujito
sujito.fmipa@um.ac.id

DOI: 10.29303/jppipa.v11i11.12428

Abstract: The quality of measurement instruments is an essential factor in ensuring accurate and meaningful learning assessments. However, many test instruments have not undergone empirical validation, which reduces their reliability and accuracy. This study aimed to develop and empirically validate a conceptual understanding test on dynamic electricity for high school students. A quantitative descriptive design was employed involving 66 students who had studied the topic of dynamic electricity. Item analysis included validity, reliability, discrimination index, and difficulty level tests. The results showed that 11 multiple-choice items met the criteria for validity and reliability with appropriate levels of difficulty and discrimination. These findings indicate that the developed instrument has good quality and can consistently measure students' conceptual understanding. Therefore, this test instrument can be used by teachers and researchers as a representative and high-quality tool to evaluate students' conceptual mastery and improve the quality of physics learning assessment.

Keywords: Conceptual understanding; Dynamic electricity; Empirical validation; Test instruments

Introduction

Understanding physics is important for supporting effective learning and learning objectives. Well-internalized concepts help students apply them to various problem contexts (Pramita & Putri, 2023; Resbiantoro et al., 2022; Susilawati et al. 2022; Wang et al., 2024). However, students frequently have difficulty understanding physics (Pribadi et al., 2023). Students have similar issues when studying dynamic electricity. Rahmawati et al. (2023) discovered a significant proportion of electrical idea misunderstandings based on the findings of a four-tier diagnostic test. High-quality instruments must be used while explaining the issues that students encounter.

The quality of instruments as tools for measuring students' conceptual understanding is an important factor in the success of assessments. The use of high-quality measuring instruments can increase engagement and participation in learning (Basri et al., 2024; Irawan, Indraloka, et al., 2025; Susilawati et al., 2022). Creating high-quality instruments requires a lengthy process and significant time (Pratama & Istiyono, 2024; Sari, 2020; Zhang et al., 2023). Another challenge in instrument development is the limitation of resources and the ability to analyze test instruments (Rismaulhijjah & Kuswanti, 2022). Basri et al. (2024) also emphasized based on their observations that the conceptual understanding test used did not meet the criteria for validity and reliability.

However, many test instruments have not undergone empirical validation, which reduces their reliability and accuracy, leading to inaccurate assessments of students' conceptual understanding and less effective learning feedback. Rigorous validation is essential because assessment users must be able to trust the results, and the proliferation of unvalidated instruments poses a significant threat to the quality and fairness of educational measurement (Kaigama & Daniel, 2022; Rezai, 2022). Considering that case,

How to Cite:
Mura, K. B., Sujito, Taufiq, A., Irawan, I. D. A., & Basri, N. A. (2025). Empirical Validation of a Conceptual Understanding Test on Dynamic Electricity for High School Students. *Jurnal Penelitian Pendidikan IPA*, *11*(11), 229–235. https://doi.org/10.29303/jppipa.v11i11.12428

research focused on developing high-quality concept understanding test instruments is needed.

Empirical validation of instruments ensures the technical quality of measurements and strengthens confidence in the resulting assessments (Camacho-Tamayo & Bernal-Ballen, 2023). In the context of physics education, instruments that have undergone empirical validation can serve as diagnostic tools, helping teachers identify students' misconceptions and design targeted learning interventions (Sukarelawan et al., 2024; Wardi & Jauhariyah, 2025). The validation process includes empirical testing using item response models or factor analysis. Through this process, items that do not represent the construct well can be revised or deleted, resulting in more accurate measurement results (Bashooir & Supahar, 2018; Kartini et al., 2019; Mardian et al., 2023). Thus, empirical validation contributes to the continuous improvement of assessment quality, ensuring that student measurement practices align with modern educational standards emphasizing evidence-based evaluation and accountability.

This study aims to test the validity of the instrument for assessing students' understanding of dynamic electricity concepts. The validity test was conducted to ensure that the resulting test instrument is of good quality and can accurately provide appropriate results (Bhakti et al., 2023; Faridah, 2021; Khasanah et al., 2019). The validated test instrument is in the form of multiple choice questions with five answer options. Multiple-choice instruments were chosen because they can present questions in a simpler and easier way (Irawan et al., 2025; Irawan et al., 2025; Kusairi, 2020). Thus, the results of this study are expected to produce a high-quality conceptual understanding test instrument that can be used by teachers and researchers in measuring students' conceptual understanding, especially in dynamic electricity.

Method

This study applied a quantitative descriptive approach with the aim of evaluating the validity of the conceptual understanding test instrument. The instrument was specifically developed to measure students' understanding of Concept of Dynamic Electricity. The research findings were used to assess the quality of the instrument as a representative measuring tool in identifying students' level of conceptual understanding. The initial instrument consisted of 13 multiple-choice questions designed based on learning outcome indicators for the topic of static electricity. The analysis results showed that out of the 13 questions tested, 11 questions met the validity criteria and were suitable for use as a conceptual understanding test. The details of the indicators for each question are presented in Table 1.

Table 1. Indicators of Conceptual Understanding Test Items

Indicator	Item No.
Analyze the relationship between current, charge, and time in an electrical circuit.	1
Apply Ohm's law in analyzing the relationship between electric current and potential difference.	2, 3, 5
Analyze resistance in an electrical circuit.	4, 6
Apply Kirchhoff's first law in investigating the current passing through a junction point.	7
Analyzing Kirchhof's second law using equations $\sum\varepsilon = \sum I \times R$	8,9
Analyzing the application of electrical power in everyday life	10
Analyzing the application of electrical energy in everyday life	11

The validation test for the conceptual understanding test instrument was conducted on 66 students who had participated in learning about dynamic electricity. The empirical validation process covered four main aspects, namely item validity testing, instrument reliability, discrimination index, and question difficulty level (Phipps et al., 2009; Rashel & Kinya, 2021; Rush et al., 2016). Prior to empirical validation, the instrument first underwent expert validation, which was conducted in January 2025. The instrument, which was determined to be eligible by the experts, was then empirically tested in February 2025 at a private secondary school in Malang City. In the implementation, students were asked to complete the test instrument independently and without reference (closed-book system) within 30 minutes.

Figure 1. Stages of empirical validation testing of conceptual understanding test instruments

The analysis consists of four stages. The first stage involves testing the validity of the instrument items. The analysis criteria are determined based on the r-calculated value obtained. If the calculated r value is greater than the critical r value, the conceptual understanding test items are declared valid. The r-table value for 66 respondents was obtained as r=0.242.

The second analysis was a reliability test of the instrument. Instruments that had been pronounced acceptable completed a reliability test to determine their consistency as measurement tools. The Cronbach Alpha value was used to determine the reliability test's conclusion criterion. A higher Cronbach Alpha indicates that an instrument is more dependable. The Cronbach Alpha score criteria are shown in Table 2.

Table 2. Reliability Test Criteria

Cronbach Alpha	Category
$0.00 \leq r < 0.20$	Unreliable
$0.20 \leq r < 0.40$	Less reliable
$0.40 \leq r < 0.60$	Fairly reliable
$0.60 \leq r < 0.80$	Reliable
$0.80 \leq r \leq 1.00$	Very reliable

The third test was conducted to measure the difficulty level of each conceptual understanding test. The test also ensured that the test items had a difficulty level appropriate to the indicators discussed. The criteria for determining the difficulty level are presented in Table 3.

Table 3. Difficulty Level Criteria

Difficulty Level	Category
$0.00 \leq D \leq 0.30$	Easy
$0.30 < D \leq 0.70$	Medium
$0.70 < D \leq 1.00$	Difficult

The next test was conducted using a discrimination test. This was done to ensure that each test item had the ability to show differences between respondents. The criteria for drawing conclusions from the discrimination test are presented in Table 4. All of the tests described above were conducted using IBM SPSS 27.

Table 4. Discrimination Index Criteria

Discrimination Index	Category
$D \leq 0.00$	Very low
$0.00 < D \leq 0.20$	Low
$0.20 < D \leq 0.40$	Medium
$0.40 < D \leq 0.70$	High
$0.70 < D \leq 1.00$	Very high

Result and Discussion

The results of the study are described based on the testing stages, including validity testing, reliability testing, discrimination index, and difficulty level. Details are provided in Table 5.

Table 5. Validation Test Results (r-table=0.242)

Item No	Pearson Correlation	Category
1	.658	Valid
2	.426	Valid
3	.506	Valid
4	.633	Valid
5	.661	Valid
6	.774	Valid
7	.386	Valid
8	.658	Valid
9	.565	Valid
10	.528	Valid
11	.593	Valid

The Pearson Correlation value is used to represent the results of the validity test of the test items for conceptual understanding. Based on a total of 66 respondents, a critical r value of 0.242 was obtained. The validity of each item was determined by comparing the calculated r value with the critical r value. As shown in Table 5, 11 items showed a calculated r value greater than the critical r value, so all of these items were declared valid in measuring students' conceptual understanding of dynamic electricity.

Table 6. Instrument reliability test

Cronbach Alpha	Category
.816	Reliable

Meanwhile, Cronbach's Alpha coefficient is used to measure the reliability of the developed instrument. The reliability value reflects the consistency and stability of the instrument in conducting measurements. An instrument is considered to have adequate reliability if its Cronbach's Alpha value is above 0.60. Based on the results shown in Table 6, the conceptual understanding test instrument for dynamic electricity material demonstrates the reliability criteria, so it can be concluded that the instrument is effective in consistently measuring students' conceptual understanding.

Table 7. Results of the Difficulty Level Test

Item No.	Items Difficult	Category
1	.697	Medium
2	.697	Medium
3	.727	Difficult
4	.545	Medium
5	.636	Medium
6	.727	Difficult
7	.485	Medium
8	.697	Medium
9	.697	Medium
10	.636	Medium
11	.515	Medium

In addition, the results of the difficulty level test also support the feasibility of the instrument. The analysis shows that two items are classified as easy, while the other nine items are classified as medium. Overall, the results of the discrimination and difficulty level analysis indicate that the conceptual understanding test instrument on concept of dynamic electricity is eligible and can be used as an effective evaluation tool in measuring students' level of conceptual understanding.

Table 8. Results of the Discrimination Index

Item No	Corrected Item	Category
1	.634	High
2	.238	Medium
3	.352	Medium
4	.564	High
5	.567	High
6	.690	High
7	.266	Medium
8	.634	High
9	.430	High
10	.470	High
11	.491	High

The discrimination index of each item in the conceptual understanding test instrument was analyzed using corrected item-total correlation scores. Based on the data presented in Table 8, eight items were classified as having high discrimination index, while the other three items were classified as having moderate discrimination index. These findings indicate that each item has a good ability to distinguish between students with high and low conceptual understanding.

Discussion

The findings of the study indicate that the dynamic electricity conceptual understanding test instrument meets the criteria for validity and is acceptable for implementation as an assessment instrument. The test instrument consists of 11 multiple-choice questions with five answer options. The questions are divided into seven indicators in the concept of dynamic electricity. In its application, students complete the instrument independently to measure their conceptual understanding. Test results are assessed by calculating the percentage of correct answers provided by the students.

The research findings also indicate that the conceptual understanding test instrument is valid based on Pearson correlation values. Eleven items were found to be valid based on the criteria with calculated r > critical r. Mujlli et al. (2023) and Ekawati et al. (2024) emphasize that a valid instrument makes an important contribution to the assessment system. These findings indicate that the developed instrument has good quality as a measurement tool for understanding. High-quality instruments can provide more accurate information from the measurements conducted (Balta & Eryılmaz, 2020; Black & Wiliam, 2009; Ntuli, 2024). An example of the developed conceptual understanding test is presented in Figure 2.

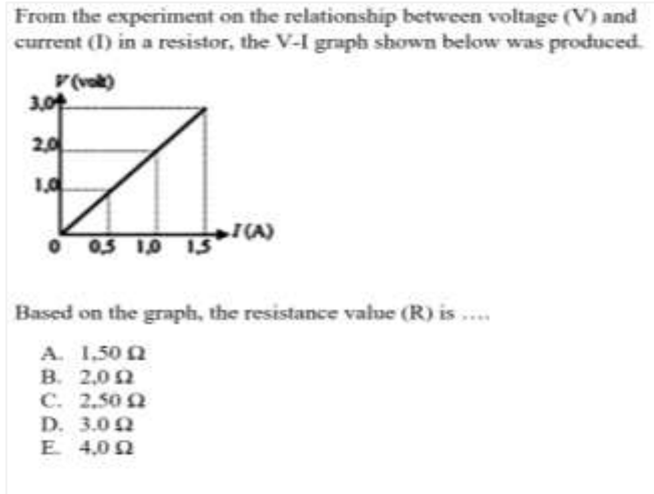

Figure 2. Display of questions on understanding dynamic electricity concepts

The reliability of the instrument also showed good results. This is because the Cronbach Alpha value obtained was in the reliable category. According to the report by Khidzir et al. (2018), instruments with a Cronbach Alpha value >0.6 can be used as test instruments. In another study, it was also reported that test instruments should have a Cronbach's alpha of more than 0.7 (Mulyana & Desnita, 2023). Reliable instruments can explain measurement results consistently in each measurement (Sukarelawan et al., 2021). In line with previous research, the assessment process must use reliable instruments (Afrizon et al., 2022).

The difficulty level of the instrument indicates that most items fall into the moderate category. According to Wulandari et al. (2022), the difficulty level can categorize students based on their ability to complete each item. These findings suggest that the instrument has an appropriate level of difficulty for measuring students' understanding of dynamic electricity concepts. Students are not overwhelmed by questions that are too difficult or too easy. This aligns with Kurniawan et al. (2024), who state that high-quality instruments have an appropriate level of difficulty. Therefore, the obtained measurement results can accurately reflect the actual outcomes (Ahmed & Ishtiaq, 2021; Danni et al., 2021).

Other findings suggest that most of the items in the instrument are highly effective in distinguishing between students with high and low abilities due to their high discrimination index. Thus, this conceptual understanding test of dynamic electricity is valid, reliable, and has good measurement sensitivity, enabling it to accurately identify students' understanding. Therefore, it is suitable for use as a measurement tool in learning evaluations because it provides meaningful information about students' achievement in dynamic electricity.

Overall, this study provides an overview of the empirical validation testing processes used to develop a quality conceptual understanding test. Teachers can use this instrument as a measurement tool when conducting classroom instruction (Sharma et al., 2021). Additionally, future researchers can use this instrument when conducting studies on related topics. However, this study has limitations, primarily the small number of respondents. Further testing with a larger sample size and respondents from diverse school backgrounds is needed to draw more accurate conclusions.

Conclusion

The study explains the results of the empirical validation testing of the conceptual understanding Test of Dynamic Electricity. It concluded that the 11 multiple-choice questions designed to measure students' understanding of dynamic electricity concepts passed the test. The test instrument has been proven to demonstratE validity and reliability criteria and have appropriate levels of difficulty and discrimination index. Therefore, future researchers can use this instrument as a representative, high-quality tool to evaluate students' conceptual understanding of dynamic electricity.

Acknowledgments
The author would like to thank the Mardi Wiyata Foundation for its support in completing this study. We also thanks to Department of Physics, State University of Malang, for their guidance and positive direction throughout the completion of this study.

Author Contributions
Karolus Boromeus Mura: Principal Investigator, Conceptualization; Sujito: Research Methodology; Ahmad Taufiq: Instrument Validation; Ivan Danar Aditya Irawan: Data Analysis; Nur Akhyar Basri: Research Implementer.

Funding
Source of funds from the Mardi Wiyata Foundation's Internal Funds.

Conflicts of Interests
The authors declare no conflict of interest.

References

Afrizon, R., Hidayati, & Novitri, L. (2022). Validity and reliability to instruments of student's initial competence for development of physics learning games integrated with scientific literacy. *Journal of Physics: Conference Series*, *2309*(1). https://doi.org/10.1088/1742-6596/2309/1/012098

Ahmed, I., & Ishtiaq, S. (2021). Reliability and Validity: Importance in medical research. *Journal of the Pakistan Medical Association*, *71*(10), 2401–2406. https://doi.org/10.47391/JPMA.06-861

Balta, N., & Eryılmaz, A. (2020). Development of Modern Physics Achievement Test: Validity and Reliability Study. *The European Educational Researcher*, *3*(1), 29–38. https://doi.org/10.31757/euer.313

Bashooir, K., & Supahar, S. (2018). Validitas dan reliabilitas instrumen asesmen kinerja literasi sains pelajaran fisika berbasis STEM. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *22*(2), 219–230. https://doi.org/10.21831/pep.v22i2.19590

Basri, N. A., Salmah, U., Irawan, I. D. A., Indraloka, R. M., & Parno, P. (2024). Rasch Analysis of the Force and Motion Conceptual Evaluation Test: Validity and Reliability in Measuring Force and Motion Understanding of Students. *Jurnal Pembelajaran Fisika*, *12*(1), 37–50. https://doi.org/10.23960/jpf.v12.n1.202404

Bhakti, Y. B., Arthur, R., & Supriyati, Y. (2023). Development of an assessment instrument for critical thinking skills in Physics: A systematic literature review. *Journal of Physics: Conference Series*, *2596*(1). https://doi.org/10.1088/1742-6596/2596/1/012067

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Camacho-Tamayo, E., & Bernal-Ballen, A. (2023). Validation of an Instrument to Measure Natural Science Teachers' Self-Perception about Implementing STEAM Approach in Pedagogical Practices. *Education Sciences*, *13*(8). https://doi.org/10.3390/educsci13080764

Danni, R., Wahyuni, A., & Tauratiya, T. (2021). Item Response Theory Approach: Kalibrasi Butir Soal Penilaian Akhir Semester Mata Pelajaran Bahasa Arab. *Arabi : Journal of Arabic Studies*, *6*(1), 93. https://doi.org/10.24865/ajas.v6i1.320

Ekawati, E. Y., Sabrilla Almas Azzahra, & Hanun Fithriyah. (2024). Construction of Critical Reasoning Skills Assessment Instruments as Diagnostic Assessments in Physics Learning with Polytomus Scoring. *Journal of Education Research and Evaluation*, *8*(2), 338–349. https://doi.org/10.23887/jere.v8i2.69259

Faridah, A. (2021). Karakteristik butir soal ujian akhir semester mata pelajaran sejarah. *Ekspose: Jurnal Penelitian Hukum dan Pendidikan, 20*(2), 1281-1288. https://doi.org/10.30863/ekspose.v20i2.1819

Irawan, I. D. A., Indraloka, R. M., Basri, N. A., Salmah, U., & Parno, P. (2025). Analysis of Concept Understanding Test Items on Static Fluid Material Using Rasch Model. *Jurnal Pendidikan Fisika, 13*(1), 1-13. https://doi.org/10.26618/jpf.v13i1.15687

Irawan, I. D. A., Basri, N. A., Kusairi, S., & Islamiyah, M. (2025). Trends and Research Directions on Formative Feedback in Physics Learning: A Systematic Review. *Jurnal Riset Pendidikan Fisika, 10*(1), 39-52. https://doi.org/10.26618/jpf.v13i2.17899

Irawan, I. D. A., Kusairi, S., Khusaini, K., Basri, N. A., Faizah, R., & Habibbulloh, M. (2025). Computer-assisted Formative Feedback: Perspectives of Physics Students and Teachers - A Comprehensive Needs Analysis. *Science Education International*, *36*(3), 311–319. https://doi.org/10.33828/sei.v36.i3.6

Kaigama, E. D., & Daniel, E. (2022). Assessment of Teachers' Attitude Towards Validation of Non-standardized Achievement Test in Secondary Schools in Borno State. *Nigeria. Education Journal*, *11*(4), 200–207. https://doi.org/10.11648/j.edu.20221104.20

Kartini, K., Doyan, A., Kosim, K., Susilawati, S., Khasanah, B. U., Hakim, S., & Muliyadi, L. (2019). Analysis of Validation Development Learning Model Attainment Concept to Improve Critical Thinking Skills and Student Learning Outcomes. *Jurnal Penelitian Pendidikan IPA, 5*(2), 185-188. https://doi.org/10.29303/jppipa.v5i2.262

Khasanah, B. U., Doyan, A., Gunawan, G., Susilawati, S., Kartini, K., Hakim, S., & Muliyadi, L. (2019). Analysis Validation of Learning Media Quantum Phenomenon. *Jurnal Penelitian Pendidikan IPA, 5*(2), 189-193. https://doi.org/10.29303/jppipa.v5i2.265

Khidzir, K. A. M., Ismail, N. Z., & Abdullah, A. R. (2018). Validity and River. *International Journal of Development and Sustainability*, *7*(3), 1026–1037. Retrieved from www.isdsnet.com/ijds

Kurniawan, A., Istiyono, E., & Naba, S. D. (2024). Item Quality Analysis of Physics Concept Understanding Test with Rasch Model. *JIPF (Jurnal Ilmu Pendidikan Fisika)*, *9*(3), 474. https://doi.org/10.26737/jipf.v9i3.5692

Kusairi, S. (2020). A web-based formative feedback system development by utilizing isomorphic multiple choice items to support physics teaching and learning. *Journal of Technology and Science Education*, *10*(1), 117–126. https://doi.org/10.3926/jotse.781

Mardian, V., Samsudin, A., Utama, J. A., Suwarma, I. R., & Coştu, B. (2023). Validity and Reliability of Elasticity Multiple-Choice Items (EMCI) Using Rasch Model. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni*, *12*(2), 265. https://doi.org/10.24042/jipfalbiruni.v12i2.17037

Mujlli, G., Al-Ghosen, A., Alrabah, R., Munshi, F., & Ozdemir, B. (2023). Development and validation of Simulation Scenario Quality Instrument (SSQI). *BMC Medical Education*, *23*(1). https://doi.org/10.1186/s12909-023-04935-5

Mulyana, V., & Desnita, D. (2023). Empirical Validity and Reliability of the Scientific Literacy Assessment Instrument Based on the Tornado Physics Enrichment Book. *Jurnal Penelitian Pendidikan IPA*, *9*(5), 3961–3967. https://doi.org/10.29303/jppipa.v9i5.3290

Ntuli, T. G. (2024). The influence of natural science scientific register in isiNdebele on classroom practices. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(4). https://doi.org/10.29333/ejmste/14417

Phipps, S. D., Brackbill, M. L., & Bernard, P. J. (2009). Relationship Between Assessment Item Format and Item Performance Characteristics. *American Journal Of Pharmaceutical Education*, *73*(9), 1–6. https://doi.org/10.5688/aj7308146

Pramita, C., & Putri, T. R. (2023). Effect of Learning Interest on Students' Concept Understanding Ability Against Subject Pressure. *International Journal of Education and Teaching Zone*, *2*(2), 243–253. https://doi.org/10.57092/ijetz.v2i2.59

Pratama, B., & Istiyono, E. (2024). Development of a Three-Tier Diagnostic Test Instrument to Identify Students' Misconceptions on Motion Kinematics Material. *Impulse: Journal of Research and Innovation in Physics Education*, *4*(2), 79–95. https://doi.org/10.14421/impulse.2023.42-02

Pribadi, F. O., Supahar, & Afriwardani, P. (2023). Certainty of Response Index-based E-Diagnostics Assisted by Google Forms to Identify Misconceptions in Simple Harmonic Waves. *Journal of Education Research and Evaluation*, *7*(1), 168–174. https://doi.org/10.23887/jere.v7i1.56351

Rahmawati, R., Widiasih, W., Marisda, D. H., & Riskawati, R. (2023). Using Four-Tier Test to Identify Prospective Elementary Teacher Students' Misconception on Electricity Topic. *Jurnal Penelitian Pendidikan IPA*, *9*(10), 7793–7802. https://doi.org/10.29303/jppipa.v9i10.3272

Rashel, U. M., & Kinya, S. (2021). Development and Validation of a Test to Measure the Secondary Students' Critical Thinking Skills: A focus on Environmental Education in Bangladesh. *International Journal of Educational Research Review*, *6*(3), 264–274. Retrieved from www.ijere.com

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

Rezai, A. (2022). Fairness in classroom assessment: development and validation of a questionnaire. *Language Testing in Asia*, *12*(1). https://doi.org/10.1186/s40468-022-00162-9

Rismaulhijjah, W. A., & Kuswanti, N. (2022). Anslisis Butir Soal Ulangan Harian Hasil Pengembangan Guru Materi Sistem Gerak Manusia Kelas XI IPA. *BIOEDU : Berkala Ilmiah Pendidikan Biologi*, *11*(3), 643–661. https://doi.org/10.26740/bioedu.v11n3.p643-661

Rush, B. R., Rankin, D. C., & White, B. J. (2016). The impact of item-writing flaws and item complexity on examination item difficulty and discrimination value. *BMC Medical Education*, *16*(1). https://doi.org/10.1186/s12909-016-0773-3

Sari, D. K. (2020). Analisis Instrumen Penilaian Kemampuan Pemodelan Matematis Pada Kelas Fisika Menggunakan Rasch Model. *MEGA : Jurnal Pendidikan Matematika*, *1*(1), 47–52. https://doi.org/10.59098/mega.v1i1.182

Sharma, U., Sokal, L., Wang, M., & Loreman, T. (2021). Measuring the use of inclusive practices among pre-service educators: A multi-national study. *Teaching and Teacher Education*, *107*, 103506. https://doi.org/10.1016/j.tate.2021.103506

Sukarelawan, M. I., Indratno, T. K., Oktova, R., & Abdullah, N. S. Y. (2024). Quality of Four-Tier Diagnostic Test on Wave and Vibration Materials: An Empirical Study Using Rasch Modeling. *Jurnal Pendidikan Fisika Indonesia*, *20*(1), 27–37. https://doi.org/10.15294/jpfi.v20i1.47493

Sukarelawan, M. I., Jumadi, Kuswanto, H., & Anas Thohir, M. (2021). The Indonesian version of the physics metacognition inventory: Confirmatory Factor Analysis and Rasch Model. *European Journal of Educational Research*, *10*(4), 2133–2144. https://doi.org/10.12973/EU-JER.10.4.2133

Susilawati, S., Doyan, A., & Muliyadi, L. (2022). Effectiveness of Guided Inquiry Learning Tools to Improve Understanding Concepts of Students on Momentum and Impulse Materials. *Jurnal Penelitian Pendidikan IPA, 8*(3), 1548-1552. https://doi.org/10.29303/jppipa.v8i3.1919

Susilawati, Doyan, A., Wahyudi, Ayub, S., & Ardhuha, J. (2022). Concept understanding of students through core physics learning tools based on guided inquiry assisted by PhET virtual media. *Journal of Physics: Conference Series*, *2165*(1). https://doi.org/10.1088/1742-6596/2165/1/012045

Wang, J., Wang, Y., Thacker, B., Wipfli, K., & Hart, S. (2024). Unveiling the influence of pedagogical content knowledge in questioning on college students' conceptual learning of introductory physics. *Physical Review Physics Education Research*, *20*(2). https://doi.org/10.1103/PhysRevPhysEducRes.20.020122

Wardi, L. Z., & Jauhariyah, M. N. R. (2025). Validity and reliability scientific literacy-based physics assessment instrument on atomic nucleus and radioactivity material. *Momentum: Physics Education Journal*, *9*(1), 109–119. https://doi.org/10.21067/mpej.v9i1.10450

Wulandari, T., Ramli, M., & Muzzazinah, M. (2022). Analisis Butir Soal Dynamic Assessment untuk Mengukur Pemahaman Konsep Klasifikasi Tumbuhan pada Mahasiswa. *Jurnal Pendidikan Sains Indonesia*, *10*(1), 191–201. https://doi.org/10.24815/jpsi.v10i1.22082

Zhang, H., Lee, I., Ali, S., DiPaola, D., Cheng, Y., & Breazeal, C. (2023). Integrating Ethics and Career Futures with Technical Learning to Promote AI Literacy for Middle School Students: An Exploratory Study. *International Journal of Artificial Intelligence in Education*, *33*(2), 290–324. https://doi.org/10.1007/s40593-022-00293-3