

# Students are Not Sure about Their Conceptual Understanding: A Comparative Study of the Level of Conceptual Understanding and the Level of Confidence Using Rasch Modeling

Toni Kus Indratno<sup>1</sup>, Moh. Irma Sukarelawan<sup>2\*</sup>, Ariati Dina Puspitasari<sup>3</sup>, Soeharto Soeharto<sup>4,5,6</sup>

<sup>1,2,3</sup> Department of Physics Education, Faculty of Teacher Training and Education, Universitas Ahmad Dahlan, Indonesia

<sup>4</sup> National Research and Innovation Agency (BRIN), Jakarta, Indonesia

<sup>5</sup> Research Center of Educational Technologies, Azerbaijan State University of Economics, Baku, Azerbaijan

<sup>6</sup> Doctoral School of Education, University of Szeged, 32-34, Petofi S. Sgt., Szeged, H-6722, Hungary

Email: [irma.sukarelawan@pfis.uad.ac.id](mailto:irma.sukarelawan@pfis.uad.ac.id)

## Article Info

### Article History

Received: Nov 30, 2022

Revised: Jan 03, 2023

Accepted: Jan 03, 2023

### Keywords:

Conceptual Understanding

Four-tier

Misconception

Rasch Model

## ABSTRACT

Misconceiving fundamental physics concepts is a persistent challenge in education, hindering students' ability to grasp more advanced topics. This study aims to compare students' conceptual understanding and confidence levels using a 4-tier diagnostic instrument. Conducted as a survey involving 56 high school students in Yogyakarta, the research utilized a 4-tier diagnostic test covering topics of temperature and heat. Data were collected through Google Forms and analyzed using Excel and Winsteps 4.6.1 based on the Rasch Model. Data analysis involved coding, preparing raw data, and utilizing Wright maps combined with Logit Value of Person (LVP) for assessing students' conceptual understanding and confidence levels. Results indicated a significant inconsistency between students' conceptual understanding and their confidence, with many students displaying high confidence despite low understanding, and vice versa. This mismatch highlights the necessity for educational strategies that not only clarify concepts but also align students' confidence with their actual understanding. The study's implications suggest the need for more comprehensive diagnostic tools and targeted pedagogical interventions to enhance learning outcomes in physics education.

This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



To cite this article:

T. K. Indratno, M. I. Sukarelawan, A. D. Puspitasari, and S. Soeharto, "Students are Not Sure about Their Conceptual Understanding: A Comparative Study of the Level of Conceptual Understanding and the Level of Confidence Using Rasch Modeling," *Indones. Rev. Phys.*, vol. 6, no. 2, pp. 75–81, 2023, doi: [10.12928/irip.v6i2.6901](https://doi.org/10.12928/irip.v6i2.6901).

## I. Introduction

Misconceptions are a classic and persistent issue in education, affecting students across various levels and disciplines, including physics [1]–[7]. These misconceptions hinder students from grasping more complex concepts, creating a snowball effect where foundational misunderstandings lead to greater difficulties in advanced topics [8]–[11]. Diagnosing and addressing these misconceptions is crucial for effective learning.

Various diagnostic instruments have been developed to identify misconceptions, ranging from open-ended questions [12], [13], to multiple-choice [11], [14], and

multi-tiered diagnostic tools [15]–[21]. Among these, the 4-Tier diagnostic instrument is particularly effective for large samples, as it separates confidence levels for conceptual questions and reasoning [22]. Despite this, there remains a lack of consensus on the best diagnostic tool, highlighting a gap in the research.

Previous studies have shown that using a combined confidence level in the three-tier method underestimates the proportion of luck and overestimates students' scores [23], [24]. However, these studies have not sufficiently explored the comparative effectiveness of the 4-Tier

instrument in diagnosing misconceptions and confidence levels.

This study aims to fill this gap by comparing the level of student's conceptual understanding and their confidence using a 4-tier diagnostic instrument. By focusing on high school students in Yogyakarta, this research will provide new insights into the alignment (or misalignment) between student's perceived and actual understanding. The findings will offer valuable contributions to educational practices and future research on diagnostic tools.

## II. Literature Review

### Misconception

Various terminologies are used to describe the differences in understanding between students and scientists. Commonly used terms are misconceptions and alternative conceptions. Misconceptions or alternative conceptions are concepts that are believed to be true but contradict the views of experts in a particular field, and show a systematic pattern of errors and are resistant to change [25]. In the context of physics, misconceptions can be interpreted as views on physics concepts that are not in line with the views of current physicists, but are believed to be true by students.

Student misconceptions result from a variety of personal experiences. The sources of these misconceptions can come from physical experience, direct observation, intuition, teaching in school and outside school, social environment, culture, language, textbooks or other teaching materials, and teachers [26].

### Four-tier diagnostic test

As the name implies, the 4-Tier type instrument consists of four levels. Level 1 (T1) is in the form of multiple-choice questions. Level 3 (T3) is the reason for T1. Level 2 (T2) is the level of confidence in T1, and Level 4 (T4) is the level of confidence in T3 [22]. The four-level type instrument is an improvement over the previous Three Tier version by adding and modifying the order of question levels. The confidence level in the three-level method is used together for the two previous levels, while the four-level method uses the confidence level for the conceptual questions and reasoning/reasoning levels separately. The use of three tiers of confidence simultaneously results in two problems: (1). Underestimating the proportion of Luck of Knowledge, and (2). Over-score students [23], [24].

### Likert Rating Scale

The Likert Rating Scale is a method often used in surveys and research to measure respondents' attitudes, opinions, or perceptions of a subject. This method involves a series of statements where respondents indicate their level of agreement or disagreement on a multi-point scale [27], [28]. The Likert scale produces ordinal data [29], which means that the intervals between points may not be equal, making it difficult to use parametric statistical methods. Therefore, caution is required in analysing Likert

scale data to avoid misinterpretation, and it is often recommended to use non-parametric statistics [30].

The variety of rating scales provides a more in-depth picture of respondents' attitudes and perceptions. In the context of misconception diagnostics, each point on this scale allows the researcher to identify the respondent's level of belief and uncertainty towards each given statement or concept. In addition to the 2-point Likert rating scale [31]–[33], 4-point [34], 5-point [35], or 6-point Likert scales [36], [37] are generally favoured in 4-tier misconception diagnostic instruments.

There are two types of Likert scales commonly used in Tier 2 or Tier 4 in four-tier diagnostic instruments, namely scales with the terms "confident" and "sure." The confident scale grades from "just guessing" (1), "very unconfident" (2), "not confident" (3), "confident" (4), "very confident" (5), to "absolutely confident" (6) [36], [38]. Meanwhile, the sure scale is graded from "I am not absolutely sure" (1), "I am not sure" (2), "I am sure" (3), and "I am absolutely sure" (4) [32]–[34].

## III. Method

This research is a survey conducted at a state senior high school in Yogyakarta, Indonesia. Fifty-six students (24 boys and 32 girls) were involved as participants or respondents. The number of respondents involved uses the assumption of the large sample size to achieve measurement accuracy and level of confidence. A sample size of 50 is used to achieve an accuracy of  $\pm 1.0$  Logit at a 99% confidence level [39]. Therefore, the involvement of 56 respondents was deemed appropriate for use. All respondents involved in this study came from class XI, who had studied heat and temperature. The "L" symbol represents the male gender, and the "P" represents the female gender.

Students' conceptual understanding of temperature and heat was carried out using a four-tier misconception diagnostic instrument (4T-HTDT) consisting of 21 items. All items are spread over four concept groups, namely: temperature (6 items), expansion (4 items), heat of change of state and temperature (4 items), and heat and displacement (7 items). Each item consists of four levels [2]. Level 1 (T1) is in the form of multiple-choice questions with several distractors. Level 2 (T2) is students' confidence in the answers they give at T1. Level 3 (T3) is the form of reasons for understanding the concepts students give to T1. Level 4 (T4) is the confidence level in students' reasons at T3. We used a 4-point Likert rating scale to explore students' confidence levels in T1 and T3. Scale 1 is used for "Just Guessing", scale 2 for "Not Sure", scale 3 for "Sure", and scale 4 for "Very Sure".

Student conceptual understanding data was collected online using 21 4T-HTDT items formatted on the Google Form platform. Google Form is distributed to students through their respective class teachers. All student participation used in this research is voluntary. Students as respondents are kept anonymous through the anonymity

model in the Google form. The demographic identity of the students involved is in the form of class and gender.

Analysis of the relationship between conceptual understanding and the level of students' confidence in the material heat and temperature using Ms. Excel and Winsteps 4.6.1 [40]. Excel is used to do the coding and prepare the raw data. At the same time, Winsteps is used to assess students' conceptual understanding and confidence levels based on the Rasch Model. Conceptual understanding and level of confidence are analyzed through a Wright map combined with Logit Value of Person (LVP) [7], [41]. Data analysis was carried out through several stages. The first stage is to carry out the process of coding the raw data. Next, the data preparation stage is carried out in \*.prn format. Then an analysis of the conceptual understanding and belief level was carried out using Winsteps. The level of conceptual understanding and the level of belief are grouped into four levels. This grouping refers to LVP through the mean and standard deviation, as shown in Table 1.

**Table 1.** Criteria for grouping the level of conceptual understanding and students' beliefs [43], [44]

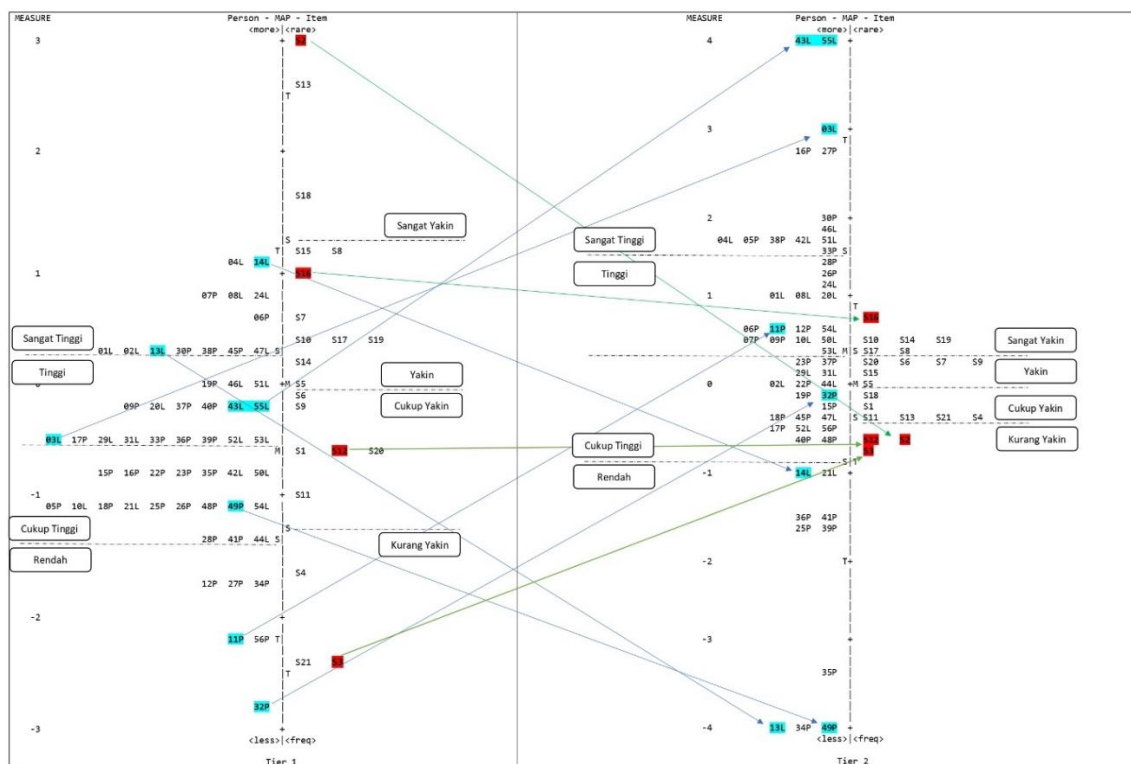
Score	Grade
$LVP < M-SD$	Low
$M-SD \leq LVP < M$	Moderate
$M \leq LVP < M+SD$	High
$LVP \geq M+SD$	Very High

#### IV. Results and Discussion

Students' concept understanding level (Tier 1 and Tier 3) and Confidence level (Tier 2 and Tier 4) are visualized using a Wright map or person-item map. A Wright map is a map that describes the state of the student's ability and the difficulty level of the item simultaneously. The Wright map is divided into two main sides: the left side, which describes the distribution of student abilities, and the right side, which describes the distribution of item difficulty levels. In principle, the distribution of student abilities and item difficulty levels is spread hierarchically from the bottom, which places the lowest logit, to the top, which places the highest logit [42]. Students with the highest ability or conceptual understanding will be placed in the top-left location, and further down for those with lower ability. At the same time, items with the highest difficulty level will be placed at the top-right, followed by items with a lower difficulty level at the bottom. On the confidence level map (Tier 2 and Tier 4), the left side visualizes gradations of students' confidence levels, and the right side visualizes gradations of questions that are believed to be correct when answered.

#### Tier 1 vs Tier 2

Comparison of students' level of understanding of concepts at Tier 1 and their level of belief at Tier 2 is shown in Figure 1.



**Figure 1.** Comparison of students' level of understanding of concepts at Tier 1 and their level of confidence at Tier 2

Figure 1 (a) shows that some students already understand the concept well. However, the average group showed that students' conceptual understanding was below

the average problem difficulty level. Thirteen of 56 people (23.2%) are in the very high category. As many as 32.1% and 33.9% are at a high and quite high level of understanding.

Only 10.7% of students have low category understanding of concepts. The S2, S13, and S18 concepts are the most difficult concepts, and the S3, S21, and S4 concepts are the easiest for students. Based on the visualization of the distribution of the difficulty level of the questions, 12 questions are in the high category (difficult and very difficult). At the same time, the other nine questions are in a low category (easy and very easy).

Figure 1 (b) shows the distribution of students' level of confidence in responding to Tier 1. Most of the students believed in the answers they gave. No less than 23.2% of students are Very Confident, and only 17.9% are Less Confident about the answers they give. Meanwhile, the percentage of students with Confident and Confident Levels were 26.8% and 32.1%, respectively. On the other hand, Figure 1(b) maps students' beliefs about the questions they answer. Three questions (S3, S2, S12) were not believed most students to answer. In comparison, the other 18 questions are well believed what they have answered. The concept most students believed to be true was the doctoral question on the concept "The temperature of a substance can be transferred". While the lowest concept believed by most students in answering was S14 regarding the concept "Heating always increases temperature".

The comparison between the level of understanding of the concept and the level of confidence held by students showed that some students needed to be more consistent between understanding the concept and the confidence they gave when answering the concept they understood. For example, students 13L and 14L are examples of

students with a very high level of understanding. Meanwhile, their level of confidence in answering is in a low category. The same thing also happened to student 49P. Student 49P has a fairly high level of understanding, but his level of confidence in his understanding is in a low category. In contrast, two students (11P and 32P) have a low level of understanding of concepts but have a high level of confidence. While three students (03L, 43L, and 55L) had a low level of understanding, they were very confident about their understanding.

A comparison of the items' difficulty level and confidence level in the items has also been carried out. The results of the analysis show that there are inconsistencies in the number of items used. For example, the S2 item "Temperature depends on the object's material" was the most difficult item for all students. However, item S2 was identified as the most trusted item. The same thing also happened to item S12 "Materials such as wool have the ability to warm the body". Contrast with item S14, "Heating always results in an increase in temperature". Item S14 is not included in the group of questions in the very difficult category but is the item that students least believe in.

### Tier 3 vs Tier 4

A comparison of students' level of understanding of concepts at Tier 3 and their level of belief at Tier 4 is shown in Figure 2.

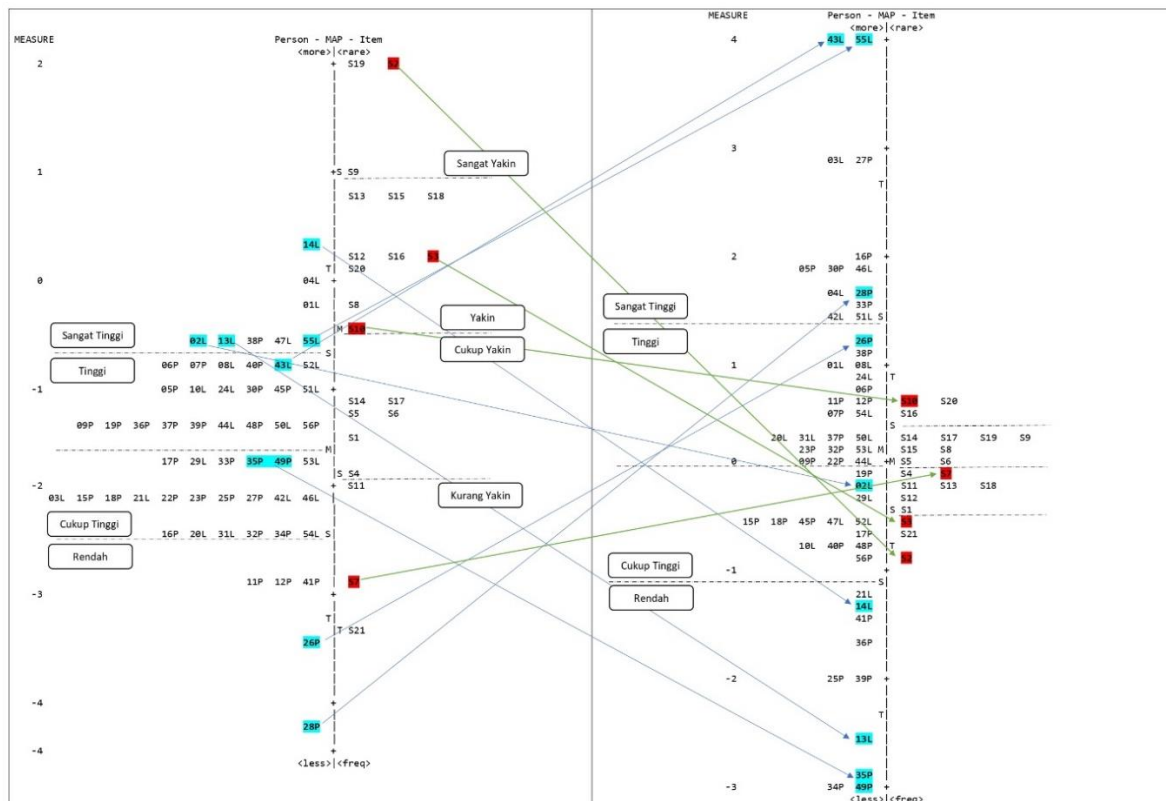


Figure 2. Comparison of students' level of understanding of concepts at Tier 3 and their level of confidence at Tier 4

Figure 2(a) shows the distribution of the correct reasons given by students for the answers they chose at Tier 1. Most students understand why the phenomenon of temperature and heat occurs at Tier 1. As many as 8 out of 56 (14.3%) students are in the Very category Tall. As many as 21 out of 56 (37.5%) students were in the high category, and 22 out of 56 (39.3%) were in the moderately high category. Only 8.9% of students are in a low category. Concepts S9, S19, and S2 are the three most difficult reasons for students to understand, and three other concepts (S11, S7, and S21) are the easiest. Twelve questions are in the high category (difficult and very difficult), and the other nine are in the low category (easy and very easy).

Figure 2(b) shows the distribution of students' level of confidence in giving reasons at Tier 3. Only a few students have low confidence in the correctness of the reasons they give. At the same time, most of them believe the reasons they give. As many as 23.2% (13 out of 56 students) really believed the reasons they gave, and only 17.9% of students were very unsure. In comparison, most of the others (58.9%) have a fairly high and high level of confidence. On the other hand, Figure 2(b) maps students' beliefs about the reasons they answered at Tier 3. A few students had low confidence in their reasons for items S10, S16, and S20. In contrast, the other three items (S2, S3, and S21) are believed to be true by most students. The S2 concept "Temperature depends on the object's material" is the most reliable truth. Meanwhile, the S10 concept "substances that expand have a fixed density." At least the truth is believed by students.

The comparison between the level of reason and the level of belief held by students shows that some students are inconsistent. Truth reasons with the beliefs they have when giving contradictory reasons. For example, student 14L is an example of one student who has a level of understanding in the highest category at Tier 3 but has a low level of confidence. In line with 14L students, 13L students are included in the category of students with very high group understanding. But has low confidence in what he has answered. The same thing also happened to students 35P and 49P. At the same time, other students have a low understanding but are very sure of the reasons they give. For example, students 28P and 26P have the lowest understanding among all students, but their confidence level is high. Identical to students 28P and 26P, students 43L and 55L have the highest confidence among all students, but their level of understanding is below that of 14L students.

Comparisons between the level of reason and the level of belief have also been carried out. The results of the analysis show that there are inconsistencies in the number of items answered. For example, the S2 item "Temperature depends on the object's material" was the most difficult item for all students. However, item S2 is identified as the item that is most believed to be the truth of the reasons given. The same thing also happens in item S3, "The temperature of a substance can be transferred". In contrast to S2 and S3, Item S10, "Substances that expand have a

constant density," is quite difficult but is at least believed by most students.

## V. Conclusion

This study highlights the importance of diagnosing students' misconceptions and their confidence levels in understanding physics concepts. Using a 4-Tier diagnostic instrument, it was found that many students have a low understanding of concepts yet show a high level of confidence in their answers, and vice versa. This inconsistency between understanding and confidence suggests that students are not always aware of their true level of understanding. Further analysis revealed that difficult items were often more believed by students than easier items. These findings emphasize the need for a pedagogical approach that focuses not only on clarifying correct concepts but also on strengthening students' beliefs and incorrect understanding. The 4-Tier diagnostic instrument proved to be effective in identifying these discrepancies, which can assist educators in designing more appropriate interventions. This study provides valuable new insights for the development of better diagnostic tools and more effective teaching strategies. In conclusion, greater efforts are needed to bridge the gap between students' conceptual understanding and beliefs to improve overall learning outcomes.

Limitations of this study include the use of a sample limited to secondary school students in Yogyakarta, so the results cannot be generalized to a wider population. In addition, the 4-Tier diagnostic instrument used has limitations in detecting the entire spectrum of misconceptions. Future research could expand the sample to different levels of education and different geographical locations to obtain more representative results. In addition, the development and testing of more comprehensive and adaptive diagnostic instruments are needed to identify misconceptions more accurately. Future research could also explore effective pedagogical interventions to correct the mismatch between concept understanding and students' belief levels. The implementation of advanced educational technologies can also be considered to improve data collection and analysis more efficiently. Thus, it is hoped that new findings can make a significant contribution to improving the quality of learning in physics and other disciplines.

## VI. Acknowledgments

We want to send our greatest gratitude to the Institute for Research and Community Service, Universitas Ahmad Dahlan.

## References

- [1] S. A. Kiray and S. Simsek, "Determination and Evaluation of the Science Teacher Candidates' Misconceptions About Density by Using Four-Tier Diagnostic Test," *Int. J. Sci.*

- Math. Educ.*, vol. 19, no. 5, pp. 935–955, Jun. 2021, doi: [10.1007/s10763-020-10087-5](https://doi.org/10.1007/s10763-020-10087-5).
- [2] M. I. Sukarelawan, S. Sriyanto, A. D. Puspitasari, D. Sulisworo, and U. N. Hikmah, “Four-Tier Heat and Temperature Diagnostic Test (4T-HTDT) to Identify Student Misconceptions,” *JIPFRI (Jurnal Inov. Pendidik. Fis. dan Ris. Ilmiah)*, vol. 5, no. 1, pp. 1–8, Jul. 2021, doi: [10.30599/jipfri.v5i1.856](https://doi.org/10.30599/jipfri.v5i1.856).
- [3] M. R. Luce and M. A. Callanan, “Family Conversations About Heat and Temperature: Implications for Children’s Learning,” *Front. Psychol.*, vol. 11, pp. 1–18, Aug. 2020, doi: [10.3389/fpsyg.2020.01718](https://doi.org/10.3389/fpsyg.2020.01718).
- [4] Z. N. Fadhilatullahifi, B. Ardiyanto, D. D. Rahayu, T. Almukholani, I. Rinayah, and F. Rahmawati, “Four-Tier Diagnostic Test Method to Identify Conceptual Understanding in Calculus,” *J. Phys. Conf. Ser.*, vol. 1613, no. 1, p. 012075, Aug. 2020, doi: [10.1088/1742-6596/1613/1/012075](https://doi.org/10.1088/1742-6596/1613/1/012075).
- [5] S. Mubarak and Y. Yahdi, “Identifying Undergraduate Students’ Misconceptions in Understanding Acid Base Materials,” *J. Pendidik. IPA Indones.*, vol. 9, no. 2, pp. 276–286, Jun. 2020, doi: [10.15294/jpii.v9i2.23193](https://doi.org/10.15294/jpii.v9i2.23193).
- [6] M. I. Sukarelawan, J. Jumadi, and N. A. Rahman, “An Analysis of Graduate Students’ Conceptual Understanding in Heat and Temperature (H&T) Using Three-Tier Diagnostic Test,” *Indones. Rev. Phys.*, vol. 2, no. 1, pp. 9–14, Jun. 2019, doi: [10.12928/irip.v2i1.910](https://doi.org/10.12928/irip.v2i1.910).
- [7] M. I. Sukarelawan, A. D. Puspitasari, D. Sulisworo, H. Kuswanto, and J. Jumadi, “A Shift in the Conceptual Understanding of Physics Students Through the Wright Map,” *J. Pendidik. Dan Pengajaran*, vol. 55, no. 1, pp. 127–141, 2022, doi: [10.23887/jpp.v55i1.38342127A](https://doi.org/10.23887/jpp.v55i1.38342127A).
- [8] Y. J. Jafer, “Assessing Kuwaiti Pre-service Science Teachers’ Greenhouse Effect Perceptions and Misconceptions,” *Int. J. Sci. Math. Educ.*, vol. 18, no. 4, pp. 657–667, Apr. 2020, doi: [10.1007/s10763-019-09992-1](https://doi.org/10.1007/s10763-019-09992-1).
- [9] G. Stylos, A. Sargioti, D. Mavridis, and K. T. Kotsis, “Validation of the thermal concept evaluation test for Greek university students’ misconceptions of thermal concepts,” *Int. J. Sci. Educ.*, vol. 43, no. 2, pp. 247–273, Jan. 2021, doi: [10.1080/09500693.2020.1865587](https://doi.org/10.1080/09500693.2020.1865587).
- [10] E. V. Slater, J. E. Morris, and D. McKinnon, “Astronomy alternative conceptions in pre-adolescent students in Western Australia,” *Int. J. Sci. Educ.*, vol. 40, no. 17, pp. 2158–2180, Nov. 2018, doi: [10.1080/09500693.2018.1522014](https://doi.org/10.1080/09500693.2018.1522014).
- [11] K. A. Bentaleb, S. Dachraoui, T. Hassouni, E. mehdi Alibrahmi, E. Chakir, and A. Belboukhari, “Development of a Survey to Assess Conceptual Understanding of Quantum Mechanics among Moroccan Undergraduates,” *Eur. J. Educ. Res.*, vol. 11, no. 4, pp. 2219–2243, Oct. 2022, doi: [10.12973/eu-jer.11.4.2219](https://doi.org/10.12973/eu-jer.11.4.2219).
- [12] A. H. Aminudin, I. Kaniawati, E. Suhendi, A. Samsudin, B. Coştu, and R. Adimayuda, “Rasch Analysis of Multitier Open-ended Light-Wave Instrument (MOLWI): Developing and Assessing Second-Years Sundanese-Scholars Alternative Conceptions,” *J. Educ. Gift. Young Sci.*, vol. 7, no. 3, pp. 557–579, Sep. 2019, doi: [10.17478/jegys.574524](https://doi.org/10.17478/jegys.574524).
- [13] I. Aykutlu, Ö. Ensari, and C. Bayrak, “Prospective teachers’ conceptual understanding of the polarization of light,” *Eur. J. Phys.*, vol. 44, no. 1, p. 015701, Jan. 2023, doi: [10.1088/1361-6404/ac93cc](https://doi.org/10.1088/1361-6404/ac93cc).
- [14] A. Hrnjičić, A. Alihodžić, F. Čunjalo, and D. Kamber Hamzić, “Development of an Item Bank for Measuring Students’ Conceptual Understanding of Real Functions,” *Eur. J. Sci. Math. Educ.*, vol. 10, no. 4, pp. 455–470, Jul. 2022, doi: [10.30935/scimath/12222](https://doi.org/10.30935/scimath/12222).
- [15] X.-M. Wang, S.-M. Wang, J.-N. Wang, G.-J. Hwang, and S. Xu, “Effects of a Two-Tier Test Strategy on Students’ Digital Game-Based Learning Performances and Flow Experience in Environmental Education,” *J. Educ. Comput. Res.*, vol. 60, no. 8, pp. 1942–1968, Jan. 2023, doi: [10.1177/07356331221095162](https://doi.org/10.1177/07356331221095162).
- [16] M. Nasir, C. Cari, W. Sunarno, and F. Rahmawati, “The Development of a Two-Tier Diagnostic Test for Student Understanding of Light,” *J. Turkish Sci. Educ.*, vol. 20, no. 1, pp. 150–172, Mar. 2023, doi: [10.36681/tused.2023.009](https://doi.org/10.36681/tused.2023.009).
- [17] W. Widodo, M. A. Mahdiannur, S. Suryanti, and N. L. Choirunnisa, “Mobile Interactive Multimedia to Assist Prospective Science Teachers Holding Conceptual Understanding in Problem-Solving Electrical Circuits,” *TEM J.*, vol. 12, no. 4, pp. 2251–2263, Nov. 2023, doi: [10.18421/TEM124-36](https://doi.org/10.18421/TEM124-36).
- [18] S. Wahyuni and T. P. Saputra, “Identification and remediation of misconceptions using three-tier diagnostic test based on CRI and literacy worksheet,” 2023, p. 050061. doi: [10.1063/5.0125856](https://doi.org/10.1063/5.0125856).
- [19] L. A. R. Laliyo, B. Sumintono, and C. Panigoro, “Measuring changes in hydrolysis concept of students taught by inquiry model: stacking and racking analysis techniques in Rasch model,” *Heliyon*, vol. 8, no. 3, p. e09126, Mar. 2022, doi: [10.1016/j.heliyon.2022.e09126](https://doi.org/10.1016/j.heliyon.2022.e09126).
- [20] J. Jumadi, M. I. Sukarelawan, and H. Kuswanto, “An investigation of item bias in the four-tier diagnostic test using Rasch model,” *Int. J. Eval. Res. Educ.*, vol. 12, no. 2, p. 622, Jun. 2023, doi: [10.11591/ijere.v12i2.22845](https://doi.org/10.11591/ijere.v12i2.22845).
- [21] N. Onder-Celikkanli and M. Tan, “Determining Turkish high school students’ misconceptions about electric charge imbalance by using a four-tier misconception test,” *Phys. Educ.*, vol. 57, no. 5, p. 055010, Sep. 2022, doi: [10.1088/1361-6552/ac68c1](https://doi.org/10.1088/1361-6552/ac68c1).
- [22] D. Kaltakci-Gurel, A. Eryilmaz, and L. C. McDermott, “Development and application of a four-tier test to assess pre-service physics teachers’ misconceptions about geometrical optics,” *Res. Sci. Technol. Educ.*, vol. 35, no. 2, pp. 238–260, Apr. 2017, doi: [10.1080/02635143.2017.1310094](https://doi.org/10.1080/02635143.2017.1310094).
- [23] D. Kaltakci Gurel, A. Eryilmaz, and L. C. McDermott, “A Review and Comparison of Diagnostic Instruments to Identify Students’ Misconceptions in Science,” *EURASIA J. Math. Sci. Technol. Educ.*, vol. 11, no. 5, pp. 989–1008, Oct. 2015, doi: [10.12973/eurasia.2015.1369a](https://doi.org/10.12973/eurasia.2015.1369a).
- [24] Soeharto, B. Csapó, E. Sarimanah, F. I. Dewi, and T. Sabri, “A Review of Students’ Common Misconceptions in Science and Their Diagnostic Assessment Tools,” *J. Pendidik. IPA Indones.*, vol. 8, no. 2, pp. 247–266, Jun. 2019, doi: [10.15294/jpii.v8i2.18649](https://doi.org/10.15294/jpii.v8i2.18649).
- [25] M. I. Sukarelawana, J. Jumadi, and H. Kuswanto, “Pengembangan Pembelajaran Mobile Remedial System (MRS) Untuk Meningkatkan Kemampuan Metakognisi dan Mereduksi Miskonsepsi Fisika pada Materi Suhu dan Kalor,” Universitas Negeri Yogyakarta, 2023.
- [26] D. Kaltakci-Gurel, A. Eryilmaz, and L. C. McDermott, “Identifying pre-service physics teachers’ misconceptions and conceptual difficulties about geometrical optics,” *Eur. J. Phys.*, vol. 37, no. 4, p. 045705, Jul. 2016, doi: [10.1088/0143-0807/37/4/045705](https://doi.org/10.1088/0143-0807/37/4/045705).
- [27] R. W. Emerson, “Likert Scales,” *J. Vis. Impair. Blind.*, vol. 111, no. 5, pp. 488–488, Sep. 2017, doi: [10.1177/0145482X1711100511](https://doi.org/10.1177/0145482X1711100511).
- [28] S. Jamieson, “Likert scales: how to (ab)use them,” *Med.*

- Educ.*, vol. 38, no. 12, pp. 1217–1218, Dec. 2004, doi: [10.1111/j.1365-2929.2004.02012.x](https://doi.org/10.1111/j.1365-2929.2004.02012.x).
- [29] S. K. Johal and M. Rhemtulla, “Comparing estimation methods for psychometric networks with ordinal data,” *Psychol. Methods*, vol. 28, no. 6, pp. 1251–1272, Dec. 2023, doi: [10.1037/met0000449](https://doi.org/10.1037/met0000449).
- [30] P. A. Bishop and R. L. Herron, “Use and Misuse of the Likert Item Responses and Other Ordinal Measures,” *Int. J. Exerc. Sci.*, vol. 8, no. 3, pp. 297–302, 2015, [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/27182418>
- [31] A. R. Suparman, E. Rohaeti, and S. Wening, “Development of Computer-Based Chemical Five-Tier Diagnostic Test Instruments: A Generalized Partial Credit Model,” *J. Effic. Responsib. Educ. Sci.*, vol. 17, no. 1, pp. 92–106, Mar. 2024, doi: [10.7160/eriesj.2024.170108](https://doi.org/10.7160/eriesj.2024.170108).
- [32] E. Istiyono, W. S. B. Dwardaru, K. Fenditasari, M. R. S. S. N. Ayub, and D. Saepuzaman, “The Development of a Four-Tier Diagnostic Test Based on Modern Test Theory in Physics Education,” *Eur. J. Educ. Res.*, vol. 12, no. 1, pp. 371–385, Jan. 2023, doi: [10.12973/eu-jer.12.1.371](https://doi.org/10.12973/eu-jer.12.1.371).
- [33] M. Irmak, H. İnaltun, J. Ercan-Dursun, H. Yanış-Kelleci, and N. Yürük, “Development and Application of a Three-Tier Diagnostic Test to Assess Pre-service Science Teachers’ Understanding on Work-Power and Energy Concepts,” *Int. J. Sci. Math. Educ.*, vol. 21, no. 1, pp. 159–185, Jan. 2023, doi: [10.1007/s10763-021-10242-6](https://doi.org/10.1007/s10763-021-10242-6).
- [34] A. C. Atmaca Aksoy and S. Erten, “A Four-Tier Diagnostic Test to Determine Pre-Service Science Teachers’ Misconception about Global Warming,” *J. Balt. Sci. Educ.*, vol. 21, no. 5, pp. 747–761, Oct. 2022, doi: [10.33225/jbse/22.21.747](https://doi.org/10.33225/jbse/22.21.747).
- [35] H. Habiddin and E. M. Page, “Uncovering Students’ Genuine Misconceptions: Evidence to Inform the Teaching of Chemical Kinetics,” *Acta Chim. Slov.*, vol. 70, no. 2, pp. 184–195, Jun. 2023, doi: [10.17344/acsi.2022.7880](https://doi.org/10.17344/acsi.2022.7880).
- [36] S. Mi, J. Ye, L. Yan, and H. Bi, “Development and validation of a conceptual survey instrument to evaluate senior high school students’ understanding of electrostatics,” *Phys. Rev. Phys. Educ. Res.*, vol. 19, no. 1, p. 010114, Feb. 2023, doi: [10.1103/PhysRevPhysEducRes.19.010114](https://doi.org/10.1103/PhysRevPhysEducRes.19.010114).
- [37] K. B. Putica, “Development and Validation of a Four-Tier Test for the Assessment of Secondary School Students’ Conceptual Understanding of Amino Acids, Proteins, and Enzymes,” *Res. Sci. Educ.*, vol. 53, no. 3, pp. 651–668, Jun. 2023, doi: [10.1007/s11165-022-10075-5](https://doi.org/10.1007/s11165-022-10075-5).
- [38] Y. K. Yan and R. Subramaniam, “Using a multi-tier diagnostic test to explore the nature of students’ alternative conceptions on reaction kinetics,” *Chem. Educ. Res. Pract.*, vol. 19, no. 1, pp. 213–226, 2018, doi: [10.1039/C7RP00143F](https://doi.org/10.1039/C7RP00143F).
- [39] B. Sumintono and W. Widhiarso, *Aplikasi model rasch untuk penelitian ilmu-ilmu sosial*. Cimahi: Trim Komunikata Publishing House, 2014.
- [40] J. M. Linacre, “Winsteps® (Version 4.6.1) [Computer Software].” 2021. [Online]. Available: <http://www.winsteps.com>
- [41] M. Ramdhan, M. I. Sukarelawan, M. A. Thohir, and F. Arifiyanti, “Junior high school student perception of online learning in pandemic Covid-19: Gender, social media ownership, and internet access duration perspective,” *Int. J. Educ. Learn.*, vol. 4, no. 1, pp. 28–40, Apr. 2022, doi: [10.31763/ijele.v4i1.517](https://doi.org/10.31763/ijele.v4i1.517).
- [42] M. I. Sukarelawan, E. Gustina, S. M. Ayu, L. Sofiana, and S. Sriyanto, “Workshop penelitian alternatif di masa pandemik covid-19 bagi guru-guru SMA/SMK,” in *Peran Perguruan Tinggi (PT) dalam Meningkatkan Kapastias Masyarakat di Era Pandemi*, 2021, pp. 1538–1544. [Online]. Available: <http://seminar.uad.ac.id/index.php/senimas/article/view/7542>
- [43] D. Adams, K. M. Chuah, B. Sumintono, and A. Mohamed, “Students’ readiness for e-learning during the COVID-19 pandemic in a South-East Asian university: a Rasch analysis,” *Asian Educ. Dev. Stud.*, vol. 11, no. 2, pp. 324–339, Mar. 2022, doi: [10.1108/AEDS-05-2020-0100](https://doi.org/10.1108/AEDS-05-2020-0100).
- [44] D. Adams, M. H. J. Tan, B. Sumintono, and S. P. Oh, “Blended Learning Engagement in Public and Private Higher Education Institutions: A Differential Item Functioning Analysis of Students’ Backgrounds,” *Malaysian J. Learn. Instr.*, vol. 17, no. 1, pp. 133–158, 2020, doi: [10.32890/mjli2020.17.1.6](https://doi.org/10.32890/mjli2020.17.1.6).

## Declarations

- Author contribution** : The authors have made significant contributions to this research. Toni Kus Indratno was responsible for designing the research and writing the manuscript as well as assisting with the data collection process. Moh. Irma Sukarelawan was responsible for analyzing the data, interpreting the findings, and helping draft the manuscript. Ariati Dina Puspitasari contributed to data collection and review. Soeharto played an important role in improving the quality of the manuscript content and improving grammar. This collaboration ensures that the research is carefully conducted and well-presented so that the results can make a meaningful contribution to the field of education.
- Funding statement** : This research was funded by Universitas Ahmad Dahlan through the Institute for Research and Community Service with the Fundamental Research (Penelitian Dasar, PD) scheme number: PD-028/SP3/LPPM-UAD/VII/2022.
- Conflict of interest** : Both authors declare that they have no competing interests.
- Additional information** : No additional information is available for this paper.