4. Quality of Four-Tier Diagnostic Test - JPFI.pdf

P-ISSN: 1693-1246
E-ISSN: 2355-3812
June 2024

Jurnal Pendidikan Fisika Indonesia 20 (1) (2024) 27-37
DOI: 10.15294/jpfi.v20i1.47493

http://journal.unnes.ac.id/nju/index.php/jpfi

Quality of Four-Tier Diagnostic Test on Wave and Vibration Materials: An Empirical Study Using Rasch Modeling

Moh. Irma Sukarelawan[1], Toni Kus Indratno[1*], Raden Oktova[1], Nurul Syafiqah Yap Abdullah[2]

[1]Physics Education Study Program,
Universitas Ahmad Dahlan, Indonesia
irma.sukarelawan@pfis.uad.ac.id, tonikus@staff.uad.ac.id, r.oktova@pfis.uad.ac.id

[3]Department of Physics,
Universiti Pendidikan Sultan Idris, Malaysia
syafiqah @fsmt.upsi.edu.my

Received: 21 February 2024. Accepted: 19 May 2024. Published: 25 June 2024

Abstract

Not all teachers have the ability to develop diagnostic instruments for misconception. Meanwhile, there is an urgent need to diagnose student misconceptions. Therefore, this research aims to carry out the process of adapting misconception-diagnostic instruments to wave and vibration materials. This study used a cross-sectional quantitative survey method. The adaptation process was conducted with 306 state high school students taking wave and vibration classes. Respondents were selected using convenience sampling techniques. 4WADI (Four-tier Wave and vibration Diagnostic Instrument) was translated into Indonesian, consulted with language experts, and assessed by material and evaluation experts to produce an instrument appropriate to the Indonesian cultural context. Before the empirical test was carried out, I-4WADI (Indonesian et al. and vibration Diagnostic Instrument) was tested for readability on 35 students. As a result, the I-4WADI has excellent readability. The validation results of six experts were analyzed using the Aiken V technique. A total of 25 items were used for expert validation. Empirical validation data were analyzed using the Rasch model. Review empirical validation from the aspects of item fit and unidimensionality. The analysis results show that I-4WADI has good quality based on validity and reliability. The infit MnSq values ranged from 0.81 to 1.40, and the outfit MnSq values ranged from 0.71 to 1.36. The instrument reliability value is 0.98. So, I-4WADI can be applied to high school students in Indonesia. These findings have practical implications for teachers who wish to diagnose students' misconceptions about waves and vibration.

Keywords: Diagnostic tests, Instrument adaptation, Misconceptions, Rasch modeling, Wave and vibration

INTRODUCTION

Understanding concepts in various fields of science is an important foundation for developing human knowledge. However, throughout the history of education, the misconception is one of the classic problems that stand out (Lemmer, Kriek, & Erasmus, 2020; Çelikkanlı & Kızılcık, 2022; Sukarelawan, Puspitasari, Sulisworo, Kuswanto, & Jumadi, 2022; Aykutlu, Ensari, & Bayrak, 2023; Brundage & Singh, 2023; Stefanou, Stylos, Georgopoulos, & Kotsis, 2023; Zhao, Zhang, Cui, Hu, & Dai, 2023). Misconception refers to a wrong or erroneous understanding of certain concepts and is believed to be accurate. Misconceptions can hinder effective learning and be detrimental to a deep understanding of the lesson material (Luce & Callanan, 2020).

Misconceptions have become a serious concern in the world of education in the last four decades, especially in the context of learning physics (Bani-Salameh, 2017; Kaniawati et al., 2019; Suhendi & Ardiansyah, 2021; Assimi, Janati Idrissi, Zerhane, & Boubih, 2022; Dwiastuti, Sukarelawan, & Sriyanto, 2022). When students develop a wrong understanding of key concepts in physics, this can interfere with their knowledge development, hinder their ability to solve problems, and even hinder their interest in studying physics.

One of the physics concepts that is important for students is the matter of wave and vibration (Somroob & Wattanakasiwich, 2017; Negoro & Karina, 2019). The concepts of wave and vibration have significantly impacted a basic understanding of the physical world and are relevant in many aspects of everyday life and other sciences. Understanding wave and vibration is the basis for understanding various other physics concepts, such as mechanics, acoustics, optics, and electromagnetism. Mobile phones, radios, computers, and medical devices such as medical imaging machines and sonar are technologies in everyday life based on the principle of wave and vibration. Understanding electromagnetic waves and vibrations helps students understand how radio, television, internet, and mobile phone signals work. Wave and vibration also have a strong connection to music and art. Understanding sound frequencies helps students understand the principles of music and acoustics. Electromagnetic waves and vibrations are also relevant to understanding the universe. For example, electromagnetic radiation from objects in space is how we study and understand galaxies, planets, and other astronomical objects. In the era of modern technology, understanding waves and vibrations is also relevant in developing technologies such as transmitting data via vibrations and electromagnetic waves (Wi-Fi, Bluetooth), facial recognition technology, and so on.

In addition to its negative impact on individual learning outcomes, misconceptions challenge teachers and educational curricula. Teachers often have to identify (Lim & Poo, 2021; Assimi, Janati, Zerhane, & Boubih ,2022; Soeharto & Csapó, 2022; Sukarelawan, Puspitasari, Sulisworo, 2022; Zhao, Zhang, Cui, Hu, & Dai 2023) and overcome misconceptions that students may have (Haidar, Yuliati, & Handayanto, 2020), while the curriculum must be carefully designed to prevent or reduce misconceptions as much as possible (Maknun & Marwiah, 2022; Syahrial, Ilmah, Yahmin, Munzil, & Muntholib, 2023). The accuracy of curriculum design to reduce and overcome misconceptions depends on diagnostic instruments as a tool for identification. Although various diagnostic instruments have been developed for this purpose, many have significant limitations. Some of these instruments may not be accurate or sensitive enough to detect certain misconceptions.

Developing a misconception diagnostic instrument is a complex task and requires special attention because it involves an in-depth understanding of the misconceptions and the ability to design effective instruments. The physics misconception diagnostic instrument must be sophisticated, relevant, and efficient to provide comprehensive information on students' understanding of concepts (Sukarelawan, Sriyanto, Puspitasari, Sulisworo, & Hikmah, 2021). In reality, not all teachers can develop misconceptions about diagnostic instruments. At the same time, there is an urgent need to use instruments to diagnose students' misconceptions.

Although experts have yet to agree, the four-level multiple choice instrument is an effective and accurate diagnostic instrument to be applied to students in large classes. Such instruments should be able to identify misconceptions appropriately, provide useful feedback to teachers and students, and be adaptable to the needs of different curricula and teaching methods.

One practical and effective instant effort is to use instruments that previous researchers have developed. However, an adaptation process is needed in different language and cultural contexts. The development of a four-level diagnostic instrument for wave and vibration materials (4WADI) was previously carried out by Caleon and Subramaniam (2010). They have reported the quality of 12 items developed in a classical test theory (CTT) perspective. By using respondents as many as 598 Singapore high school students, the 4WADI instrument has shown its performance in diagnosing student misconceptions. However, 4WADI cannot be applied in the cultural context of Indonesia because it requires an adaptation process. In addition, there is limited information that can be reported when using CTT (Sukarelawan, Jumadi, Kuswanto, Soeharto, & Hikmah, 2021; Guo, Lu, Johnson, & McCaffrey, 2022). The difficulty level of items in CTT depends on the sample or the respondent's ability (Caycho-Rodríguez et al., 2022). This results in a person's score being low if the question is difficult and high if the question is easy. Therefore, other approaches are needed, such as Rasch modeling based on modern test theory, which can be an alternative to CTT. Rasch modeling can be an integrated approach to provide more comprehensive information regarding the psychometric properties of 4WADI that have not yet been reported (Sukarelawan, Jumadi, Kuswanto, & Thohir, 2021). Through Rasch modeling, the psychometric properties of 4WADI in Indonesian culture can be evaluated at the individual item level. The use of Rasch modeling to evaluate the psychometric properties of 4WADI in various cultures has not been widely reported, including in Indonesia. Therefore, this research aims to develop a diagnostic instrument for misconceptions regarding waves and vibrations through an adaptation process in the Indonesian cultural context.

METHOD

This study used a cross-sectional quantitative survey method. The main sample used in this study was 306 public high school students selected using the convenience sampling technique. Respondents consisted of 67.6% (207 of 306 students) female and 32.4% (99 of 306 students) male. All respondents came from classes XI and XII from 3 public schools. The sample size used has considered the measurement accuracy to be 0.5 logit accuracy and a 99% confidence level (Chen et al., 2014; Lee, Chinna, & Sumintono, 2020).

The main instrument of this research is the four-tier wave and vibration diagnostic instrument for misconception (4WADI), which consists of 12 items. As the name suggests, each question in the diagnostic test instrument has four levels of questions. The first level is conceptual questions. The second level is the respondent's level of confidence in answering the first level. The third level is a choice of explanation of the phenomenon at the first level. The fourth level is the level of confidence in the choices given at the third level. The cognitive level of the questions used is C2 (understanding) to C4 (analyzing). The supporting instruments used to support the adaptation process in the Indonesian cultural context include advice sheets from language experts, expert assessment sheets, and readability test sheets for students. The language expert's suggestion sheet contains a column of suggestions and improvements descriptively. The expert assessment sheet comprises 25 items spread across material, construction, and language/cultural aspects. The readability test sheet for students consists of 12 items spread over aspects of work instructions, content aspects, linguistic aspects, and graphical aspects.

The initial procedure was to adapt the instrument to the Indonesian cultural context. The researcher translated the language and then consulted with a linguist to ensure equality between the original and Indonesian versions (I-4WADI). After going through the adaptation process, I-4WADI was reassessed by six material and evaluation experts. After being declared suitable by experts, I-4WADI was tested for readability on 35 students who had taken the wave and vibration class. Readability tests were analyzed using scoring techniques. After passing the readability test, I-4WADI was tested empirically on students. Empirical tests are carried out online to facilitate the data administration process. Students complete the diagnostic process in between 30 to 40 minutes.

The data obtained in this study are in the form of suggestions and corrections from language experts in the form of qualitative data. Expert validation data were analyzed using the Aiken V technique. Data from the readability test results were analyzed using a scale conversion guideline of 4. The psychometric properties of I-4WADI were empirically analyzed using Rasch modeling. The reported index is in the form of reliability in terms of separation and reliability aspects (Yuhanna, Al-Muhdhar, Gofur, & Hassan, 2021). Then, use the Logit Value of the Item (LVI) combined with the Wright map to evaluate the distribution of the difficulty levels of the questions (Sukarelawan et al., 2021; Sukarelawan et al. et al., 2022). The fit analysis to the Rasch modeling was evaluated based on the infit and outfit MnSq (Matheny & Clanton, 2020). The acceptance limit for the model's fit is when the MnSq infit and outfit values are in the range of 0.5–1.5 (Sumintono & Widhiarso, 2014, 2015). Another index reported is unidimensionality. I-4WADI is unidimensional if the score of raw variances explained by measure exceeds 30%, and the Unexplained variance in 1st contrast does not exceed 15%.

RESULT AND DISCUSSION

1. Content Validity

Before being validated by experts, a four-tier diagnostic instrument on wave and vibration material (4WADI) was adapted using the model proposed by Muñiz, Elosua, Padilla, and Hambleton (2016). A summary of content validity is shown in Figure 1.

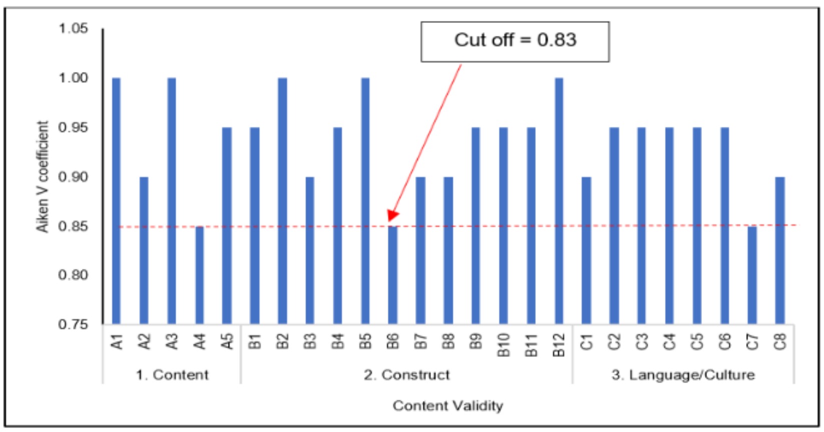

Figure 1. Aiken V validity coefficient

Content validity plays an important role in developing tests to achieve predictive validity. Validation results from a number of material experts and evaluation experts show that the Aiken V coefficient value is in a very satisfactory range, namely between 0.85 and 1.00. This indicates that the diagnostic instrument used in this research has a very high level of validity. In the context of the minimum limit value of the validity criteria used, all items in the misconception diagnostic instrument met the established standards of 0.83 (Larasati, Supahar, & Yunanta, 2020). Therefore, this instrument can be relied on to identify misconceptions in wave and vibration materials. These findings align with the results of the content validity test carried out by (Caleon & Subramaniam, 2010).

2. Respondent readability test results

After being translated into I-4WADI and declared valid by experts, I-4WADI was tested for readability on 35 students. The readability test is an important process in I-4WADI adaptation. The readability test helps measure the extent to which students can read and understand the instrument easily. By identifying potential ambiguity, confusion, or difficult language, readability tests can help improve the accuracy of I-4WADI.

Readability evaluation refers to four important aspects: work instructions, content, language, and graphics. The results of this assessment are very satisfying. First, in the aspect of work instructions, this diagnostic instrument obtained an average score of 3.37. This shows that the instructions given to students are very good and easy to understand. This assessment falls into the "Very Good" category, indicating that the instrument provides clear directions for students in carrying out their assignments.

Before entering the diagnostic section, the initial part of the I-4WADI consists of instructions for use and respondent demographics. The information contained in the I-4WADI consists of filling requirements and a brief explanation of the question model used in the instrument. The instructions section of the misconception diagnostic instrument is very important for potential respondents because it provides clear directions on how to answer the questions correctly. Good instructions can help potential respondents understand the purpose of the instrument and how they should interpret the questions and provide accurate responses. This can also help reduce misunderstandings or wrong interpretations that can interfere with the validity of measurement results.

Second, in terms of content, the average value reached 3.45. This shows that the material presented on the diagnostic instrument is very relevant and comprehensive. This assessment is also included in the "Very Good" category, which shows that the content of this instrument really supports students' understanding of waves and vibrations. This cannot be separated from the quality of 4WADI, as Caleon and Subramaniam (2010) reported. They have undergone a series of structured and systematic steps in producing 4WADI, starting from Content Validation and Piloting, Construction, Administration, and Validation.

Third, in the linguistic aspect, I-4WADI obtained an average score of 3.43, which is also included in the "Very Good" category. This shows that the language used is very precise and easy for students to understand. Linguists have guaranteed equivalence between the original language version of the instrument (4-WADI) and the Indonesian version (I-4WADI) through a forward-backward translation procedure. They are responsible for ensuring that the translation reflects the original text's exact meaning and takes into account the cultural and linguistic context of the target language. Linguist experts provide recommendations and ensure the target Indonesian students understand the translated instruments well. More than that, linguists also recommend some relevant vocabulary, avoiding phrases or terms that are ambiguous or difficult for potential respondents to understand.

Finally, the graphic aspect, which is also important in diagnostic instruments, has an average value of 3.37. Like the three previous aspects, this graphic aspect is also included in the "Very Good" category. This shows that the appearance and arrangement of the instruments are very good and support students' experience in answering questions.

Thus, the readability results of the four-level diagnostic instrument on wave and vibration materials show that it is in the "Excellent" category in all evaluated aspects. This is a very positive indication that this instrument is effective in helping students understand the questions used in the diagnostic instrument. The readability test results of I-4WADI in the Indonesian cultural context are parallel to the readability test results carried out on the original version (Caleon & Subramaniam, 2010).

3. Empirical test results

a. Item reliability

The reliability analysis results of the instrument's items are displayed in a summary, as shown in Table 1. Item reliability is a measure that indicates the extent to which the items in the test are consistent in measuring individual abilities (Zmnako & Chalabi, 2019).

Table 1. Summary of I-4WADI statistics

Index of item	Score
Infit MnSq	1.01
Infit ZStd	-0.08
Outfit MnSq	0.99
Outfit ZStd	-0.26
Separation	6.46
Reliability	0.98

The item reliability value in Table 1 is 0.98, indicating that the items in the test have a very high level of reliability in measuring individual abilities (Monaco et al., 2022). This is a positive indicator that the items are reliable in measuring the desired misconceptions. This shows that the items in the test consistently measure the ability to understand individual concepts. The reliability at the item level complements the reliability of the original 4WADI, which has yet to be reported by Caleon and Subramaniam (2010).

Item separation is a measure that indicates the extent to which the items in a test can differentiate between individuals with different abilities (Cordier et al., 2018). The item separation value of 6.46 indicates that the items in the test have a very good ability to differentiate between individuals with different abilities (Adams, Tan, & Sumintono, 2021). The higher the value of the separation item, the better the items in the test are in discriminating between individuals with different abilities.

b. Item Difficulty Distribution (LVI and Wright map)

The distribution of items at each difficulty level is visualized in Figure 2. Figure 2, the Person-item map, shows the two items (Q9 and Q12) that were easiest for students. The adequacy of the number of items, the distribution hierarchy, and the existence of gaps in the item difficulty hierarchy are indicators that support the representativeness of the content (Ha, 2021). Figure 2 shows the sufficient number of items and their distribution in each difficulty hierarchy. However, there is a significant gap between items Q12 and Q9. From the person-item map in Figure 2, there is space for item improvement by adding items that are more difficult to answer to cover all person abilities.

c. Item fit

To obtain information on item suitability for the Rasch modeling, the infit and outfit MnSq values were determined. Table 2 details the infit and outfit MnSq values of each item. The infit and outfit MnSq values provide information about the item's fit to the Rasch model. In this study, the infit MnSq values ranged from 0.81 to 1.40, and the outfit MnSq values ranged from 0.71 to 1.36. These values indicate that the items fit the Rasch model. Values in the range of 0.5 to 1.5 are generally considered acceptable (Li et al., 2023; Palacios-Espinosa et al., 2023; Sumintono, 2015).

The outfit Zstd value, which measures standard residuals, ranges from -3.03 to 3.65. These values indicate the extent to which the observed response deviates from the expected response based on the Rasch model. The range of values indicates that some items may not fit the model because values outside the range of -2 to 2 are usually considered problematic (Rusmansyah & Almubarak, 2020).

d. Unidimensionality

Table 2 shows the unidimensionality of the items used. Residual Principal Component Analysis has been used to determine the unidimensionality of instruments (Han, 2022; Liu & Boone, 2023). This unidimensionality indicates that the instrument measures one single dimension in the latent factors investigated in this study (Scoulas, Aksu Dunya, & De Groote, 2021; Simpelaere, Van Nuffelen, De Bodt, Vanderwegen, & Hansen, 2017). For the instrument developed to be valid in measuring latent factors or unidimensionality, it must meet the standard that the score of raw variances explained by measure exceeds 30% (Chou & Wang, 2010). However, the diagnostic instrument developed had a score of raw variances explained by the measure of 28.4%. Even though it is below the 30% limit, this figure shows that the unidimensionality of the instrument being developed remains relatively good.

```
 MEASURE    Person - MAP - Item
                <more>|<rare>
    1               .  +T
                    .  |
                    #  |
                    .  |  Q11
                    #  |  Q1
                    .  |S Q5
                    .  |  Q4
                    . T|
                       |  Q8
                    #  |
                       |M Q2      Q7
                 ####  |  Q10     Q6
    0             .## S+  Q3
                .####  |
                 .###  |
                .####  |S Q9
               ######  |
              .###### M|
         #############  |
             .########  |
               #######  |T
                       |
                 .### S|
                  .##  |  Q12
                       |
   -1            ####  +
                       |
                      T|
                    .  |
                       |
                       |
                    .  |
                       |
                       |
                       |
                       |
                       |
   -2               .  +
                <less>|<freq>
 EACH "#" IS 4: EACH "." IS 1 TO 3
```

Figure 2. Distribution of items in the person-item map (Wright map)

Table 2. The fit of Items to Rasch Models

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
1	743	306	.49	.07	1.40	3.03	1.25	1.59	.13	.35	62.7	64.6	Q1
2	878	306	-.04	.06	1.07	.99	1.10	1.13	.46	.39	19.0	25.2	Q2
3	931	306	-.19	.05	.83	-3.02	.83	-2.36	.43	.40	26.5	22.9	Q3
4	787	306	.28	.06	1.12	1.21	1.16	1.29	.10	.37	52.3	53.4	Q4
5	771	306	.35	.07	.94	-.57	.87	-.96	.44	.36	53.6	54.4	Q5
6	885	306	-.06	.05	1.05	.80	1.03	.38	.45	.39	19.6	25.3	Q6
7	860	306	.02	.06	.76	-3.53	.74	-3.03	.54	.39	33.7	30.5	Q7
8	828	306	.12	.06	.81	-2.39	.71	-3.01	.58	.38	41.8	37.3	Q8
9	1020	306	-.42	.05	1.03	.52	.98	-.30	.43	.40	16.0	20.7	Q9
10	891	306	-.08	.05	.90	-1.52	.89	-1.30	.35	.39	21.6	25.3	Q10
11	730	306	.56	.08	.88	-.95	.96	-.21	.29	.34	65.4	68.5	Q11
12	1252	306	-1.03	.05	1.34	4.45	1.36	3.65	.23	.35	22.2	30.8	Q12
MEAN	881.3	306.0	.00	.06	1.01	-.1	.99	-.3			36.2	38.2	
P.SD	137.1	.0	.41	.01	.19	2.3	.19	1.9			17.4	16.5	

Table 3. Unidimensionality of items

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units

		Eigenvalue	Observed		Expected
Total raw variance in observations	=	16.7485	100.0%		100.0%
Raw variance explained by measures	=	4.7485	28.4%		28.4%
Raw variance explained by persons	=	.5610	3.3%		3.4%
Raw Variance explained by items	=	4.1875	25.0%		25.1%
Raw unexplained variance (total)	=	12.0000	71.6%	100.0%	71.6%
Unexplned variance in 1st contrast	=	2.0762	12.4%	17.3%	
Unexplned variance in 2nd contrast	=	1.3695	8.2%	11.4%	
Unexplned variance in 3rd contrast	=	1.2407	7.4%	10.3%	
Unexplned variance in 4th contrast	=	1.1866	7.1%	9.9%	
Unexplned variance in 5th contrast	=	1.1818	7.1%	9.8%	

CONCLUSION

A four-level diagnostic instrument on wave and vibration materials has been developed through an adaptation process in the Indonesian cultural context. The results of adapting 4WADI to the Indonesian cultural context have been carried out into I-4WADI, and the results show that the four-level type diagnostic instrument on wave and vibration material is appropriate to the cultural context of high school students in Indonesia—the results of the expert assessment of content validity range from 0.85 to 1.00. The readability level is very good, with an average score of 3.37. Empirically, I-4WADI has infit MnSq values ranging from 0.81 to 1.40, and the outfit MnSq values ranged from 0.71 to 1.36. The instrument reliability value is 0.98. Therefore, I-4WADI meets the criteria for validity and reliability.

This research has the potential to make a significant contribution to curriculum development and improve the quality of education in Indonesia through the instrument adaptation process that has been carried out. However, these findings cannot be generalized to a wider context because the heterogeneity of respondents is limited to public high school students. In future research, the heterogeneity of respondents from private and vocational schools can be considered by future researchers so that it will provide a richer and more comprehensive view.

ACKNOWLEDGMENTS

We want to send our greatest gratitude to the Directorate of Research and Community Service, Ministry of Education, Culture, Research and Technology of the Republic of Indonesia for granting research funds number: 015/PFR/LPPM UAD/VI/2023 in the Fundamental Research scheme.

REFERENCES

Adams, D., Tan, M. H. J., & Sumintono, B. (2021). Students' readiness for blended learning in a leading Malaysian private higher education institution. *Interactive Technology and Smart Education*, 18(4), 515–534. Retrieved from https://doi.org/10.1108/ITSE-03-2020-0032

Assimi, E., Janati I, R., Zerhane, R., & Boubih, S. (2022). The use of a three-tier diagnostic test to investigate conceptions related to cell biology concepts among pre-service teachers of life and earth sciences. *Journal of Biological Education*, 1–28. Retrieved from https://doi.org/10.1080/00219266.2022.2134175

Aykutlu, I., Ensari, Ö., & Bayrak, C. (2023). Prospective teachers' conceptual understanding of the polarization of light. *European Journal of Physics*, 44(1), 015701. Retrieved from https://doi.org/10.1088/1361-6404/ac93cc

Bani-Salameh, H. N. (2017). How persistent are the misconceptions about force and motion held by college students? *Physics Education*, 52(1), 14003.

Brundage, M. J., & Singh, C. (2023). Development and validation of a conceptual multiple-choice survey instrument to assess student understanding of introductory thermodynamics. *Physical Review Physics Education Research*, 19(2), 020112. Retrieved from https://doi.org/10.1103/PhysRevPhysEducRes.19.020112

Caleon, I. S., & Subramaniam, R. (2010). Do Students Know What They Know and What They Don't Know? Using a Four-Tier Diagnostic Test to Assess the Nature of Students' Alternative Conceptions. *Research in Science Education*, 40(3), 313–337. Retrieved from https://doi.org/10.1007/s11165-009-9122-4

Caycho-Rodríguez, T., Vilca, L. W., Cervigni, M., Gallegos, M., Martino, P., Portillo, N., & Burgos Videla, C. (2022). Fear of COVID-19 scale: Validity, reliability and factorial invariance in Argentina's general population. *Death Studies*, *46*(3), 543–552. doi:10.1080/07481187.2020.1836071

Çelikkanlı, N. Ö., & Kızılcık, H. Ş. (2022). A Review of Studies About Four-Tier Diagnostic Tests in Physics Education. *Journal of Turkish Science Education*, 19(4), 1291–1311. Retrieved from https://doi.org/10.36681/tused.2022.175

Chen, W.-H., Lenderking, W., Jin, Y., Wyrwich, K. W., Gelhorn, H., & Revicki, D. A. (2014). Is Rasch model analysis applicable in small sample size pilot studies for assessing item characteristics? An example using PROMIS pain behavior item bank data. *Quality of Life Research*, 23(2), 485–493. Retrieved from https://doi.org/10.1007/s11136-013-0487-5

Chou, Y.-T., & Wang, W.-C. (2010). Checking Dimensionality in Item Response Models with Principal Component Analysis on Standardized Residuals. *Educational and Psychological Measurement*, 70(5), 717–731. Retrieved from https://doi.org/10.1177/0013164410379322

Cordier, R., Speyer, R., Schindler, A., Michou, E., Heijnen, B. J., Baijens, L., & Joosten, A. V. (2018). Using Rasch Analysis to Evaluate the Reliability and Validity of the Swallowing Quality of Life Questionnaire: An Item Response Theory Approach. *Dysphagia*, 33(4), 441–456. Retrieved from https://doi.org/10.1007/s00455-017-9873-4

Dwiastuti, E., Sukarelawan, M. I., & Sriyanto, S. (2022). Learning Management System Acceptance Rate Among Vocational School Physics' Student. *JIPFRI (Jurnal Inovasi Pendidikan Fisika Dan Riset Ilmiah)*, 6(1), 10–18. Retrieved from https://doi.org/10.30599/jipfri.v6i1.1183

Guo, H., Lu, R., Johnson, M. S., & McCaffrey, D. F. (2022). Alternative Methods for Item Parameter Estimation: From CTT to IRT. *ETS Research Report Series*, 2022(1), 1–16. Retrieved from https://doi.org/10.1002/ets2.12355

Ha, H. T. (2021). A Rasch-based validation of the Vietnamese version of the Listening Vocabulary Levels Test. *Language Testing in Asia*, 11(1), 16. Retrieved from https://doi.org/10.1186/s40468-021-00132-7

Haidar, D. A., Yuliati, L., & Handayanto, S. K. (2020). The effect of inquiry learning with scaffolding on misconception of light material among fourth-grade students. *Jurnal Pendidikan IPA Indonesia*, 9(4), 540–553. Retrieved from https://doi.org/10.15294/jpii.v9i4.22973

Han, H. (2022). The effectiveness of weighted least squares means and variance adjusted based fit indices in assessing local dependence of the rasch model: Comparison with principal component analysis of residuals. *PLOS ONE*, 17(9), e0271992. Retrieved from https://doi.org/10.1371/journal.pone.0271992

Kaniawati, I., Fratiwi, N. J., Danawan, A., Suyana, I., Samsudin, A., & Suhendi, E. (2019). Analyzing students' misconceptions about Newton's Laws through Four-Tier Newtonian Test (FTNT). *Journal of Turkish Science Education*, 16(1), 110–122. Retrieved from https://doi.org/10.12973/tused.10269a

Larasati, P. E., Supahar, & Yunanta, D. R. A. (2020). Validity and reliability estimation of assessment ability instrument for data literacy on high school physics material. *Journal of Physics: Conference Series*, 1440(1), 012020. Retrieved from https://doi.org/10.1088/1742-6596/1440/1/012020

Lemmer, M., Kriek, J., & Erasmus, B. (2020). Analysis of Students' Conceptions of Basic Magnetism from a Complex Systems Perspective. *Research in Science Education*, 50(2), 375–392. Retrieved from https://doi.org/10.1007/s11165-018-9693-z

Li, D., Zong, X., Huang, Q., Wu, F., Huang, Y., Ge, Y., … Yuan, C. (2023). Validation of the simplified Chinese version of PROMIS Parent Proxy-25 Profile in parents of children with cancer. *Journal of Pediatric Nursing*, in press. Retrieved from https://doi.org/10.1016/j.pedn.2023.05.016

Lim, H. L., & Poo, Y. P. (2021). Diagnostic test to assess misconceptions on photosynthesis and plant respiration: Is it valid and reliable? *Jurnal Pendidikan IPA Indonesia*, 10(2), 241–252. Retrieved from https://doi.org/10.15294/jpii.v10i2.26944

Ling Lee, W., Chinna, K., & Sumintono, B. (2020). Psychometrics assessment of HeartQoL questionnaire: A rasch analysis. *European Journal of Preventive Cardiology*, 2047487320902322. Retrieved from https://doi.org/10.1177/2047487320902322

Liu, X., & Boone, W. J. (2023). Introduction to Advances in Applications of Rasch Measurement in Science Education. In X. Liu & W. J. Boone (Eds.), *Advances in Applications of Rasch Measurement in Science Education. Contemporary Trends and Issues in Science Education* (pp. 1–18). Springer. Retrieved from https://doi.org/10.1007/978-3-031-28776-3_1

Luce, M. R., & Callanan, M. A. (2020). Family Conversations About Heat and Temperature: Implications for Children's Learning. *Frontiers in Psychology*, 11(August), 1–18. Retrieved from https://doi.org/10.3389/fpsyg.2020.01718

Maknun, J., & Marwiah, M. (2022). Remediation of Misconceptions Vocational High School Students on the Concept of Static Fluids using the Conceptual Change Model. *Journal of Technical Education and Training*, 14(2). Retrieved from https://doi.org/10.30880/jtet.2022.14.02.005

Matheny, L. M., & Clanton, T. O. (2020). Rasch Analysis of Reliability and Validity of Scores From the Foot and Ankle Ability Measure (FAAM). *Foot and Ankle International*, 41(2), 229–236. Retrieved from https://doi.org/10.1177/1071100719884554

Monaco, E., Pisanu, G., Carrozzo, A., Giuliani, A., Conteduca, J., Oliviero, M. & Ferretti, A. (2022). Translation, cross-cultural adaptation, and validation of the Italian version of the anterior cruciate ligament–return to sport after injury (ACL-RSI) scale and its integration into the K-STARTS test. *Journal of Orthopaedics and Traumatology*, 23(1), 1–8. Retrieved from https://doi.org/10.1186/s10195-021-00622-7

Muñiz, J., Elosua, P., Padilla, J.-L., & Hambleton, R. K. (2016). Test Adaptation Standards for Cross-Lingual Assessment. In C. S. Wells, M. Faulkner-Bond, & E. Hambleton (Eds.), *Educational Measurement: From Foundations to Future* (pp. 291–304). The Guilford Press.

Negoro, R. A., & Karina, V. (2019). Development of a Four-Tier Diagnostic Test for Misconception of Oscillation and Waves. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 5(2), 69–76. Retrieved from https://doi.org/10.21009/1.05201

Palacios-Espinosa, X., Pedraza, R. S., Gómez-Carvajal, A.-M., Botero-Meneses, J. S., Escallón, D. M., & Leal, D. A. (2023). Psychometric Properties of the Death Anxiety Scale for Adult Chronic Patients. *Trends in Psychiatry and Psychotherapy*, in press. Retrieved from https://doi.org/10.47626/2237-6089-2023-0630

Rusmansyah, R., & Almubarak, A. (2020). Students' Cognitive Analysis using Rasch Modeling as an Assessment for Planning of Strategies in Chemistry Learning. *JKPK (Jurnal Kimia Dan Pendidikan Kimia)*, 5(3), 222. Retrieved from https://doi.org/10.20961/jkpk.v5i3.40936

Scoulas, J. M., Aksu Dunya, B., & De Groote, S. L. (2021). Validating students' library experience survey using rasch model. *Library & Information Science Research*, 43(1), 101071. Retrieved from https://doi.org/10.1016/j.lisr.2021.101071

Simpelaere, I. S., Van Nuffelen, G., De Bodt, M., Vanderwegen, J., & Hansen, T. (2017). Validation of the Dutch version of the Swallowing Quality-of-Life Questionnaire (DSWAL-QoL) and the adjusted DSWAL-QoL (aDSWAL-QoL) using item analysis with the Rasch model: a pilot study. *Health and Quality of Life Outcomes*, 15(1), 66. Retrieved from https://doi.org/10.1186/s12955-017-0639-3

Soeharto, S., & Csapó, B. (2022). Exploring Indonesian Student Misconceptions in Science Concepts. *Heliyon*, 8(9), e10720. Retrieved from https://doi.org/10.1016/j.heliyon.2022.e10720

Somroob, S., & Wattanakasiwich, P. (2017). Investigating student understanding of simple harmonic motion. *Journal of Physics: Conference Series*, 901, 012123. Retrieved from https://doi.org/10.1088/1742-6596/901/1/012123

Stefanou, M., Stylos, G., Georgopoulos, K., & Kotsis, K. (2023). Primary Preservice Teachers' Misconceptions and Reasoning of Thermal Concepts in Everyday Contexts. *The International Journal of Learning in Higher Education*, 31(1), 127–157. Retrieved from https://doi.org/10.18848/2327-7955/CGP/v31i01/127-157

Suhendi, H. Y., & Ardiansyah, R. (2021). Development of HATRADI : a four tier test for diagnostic student misconception in heat transfer concept. *Journal of Physics: Conference Series*, 1918(2). Retrieved from https://doi.org/10.1088/1742-6596/1918/2/022015

Sukarelawan, M. I., Jumadi, J., Kuswanto, H., Soeharto, S., & Hikmah, F. N. (2021). Rasch Analysis to Evaluate the Psychometric Properties of Junior Metacognitive Awareness Inventory in the Indonesian Context. *Jurnal Pendidikan IPA Indonesia*, 10(4), 486–495. Retrieved from https://doi.org/10.15294/jpii.v10i4.27114

Sukarelawan, M. I., Jumadi, J., Kuswanto, H., & Thohir, M. A. (2021). The Indonesian Version of the Physics Metacognition Inventory: Confirmatory Factor Analysis and Rasch Model. *European Journal of Educational Research*, *10*(4), 2133–2144. doi:10.12973/eu-jer.10.4.2133

Sukarelawan, M. I., Puspitasari, A. D., Rahmatika, Z. 'Amalia, Dennis, D., Ishafit, I., Hikmah, F. N., & Sulistyo, E. N. (2022). Online Learning Using Google Incorporated for Student High School: Mapping Motivation Using Rasch Model in Physics Learning. *Berkala Ilmiah Pendidikan Fisika*, 10(2), 216–224. Retrieved from https://doi.org/10.20527/bipf.v10i2.12209

Sukarelawan, M. I., Puspitasari, A. D., Sulisworo, D., Kuswanto, H., & Jumadi. (2022). A shift in the conceptual understanding of physics students through the wright map. *Jurnal Pendidikan Dan Pengajaran*, 55(1), 127–141. Retrieved from https://doi.org/10.23887/jpp.v55i1.37155

Sukarelawan, M. I., Sriyanto, S., Puspitasari, A. D., Sulisworo, D., & Hikmah, U. N. (2021). Four-Tier Heat and Temperature Diagnostic Test (4T-HTDT) to Identify Student Misconceptions. *JIPFRI (Jurnal Inovasi Pendidikan Fisika Dan Riset Ilmiah)*, 5(1), 1–8. Retrieved from https://doi.org/10.30599/jipfri.v5i1.856

Sumintono, B. (2015). Pemodelan Rasch pada Asesmen Pendidikan: Suatu Pengantar. *Konferensi Guru Dan Dosen Nasional (KGDN) 2015*, (November), 1–14. Retrieved from http://eprints.um.edu.my/id/eprint/15603%0A

Sumintono, B., & Widhiarso, W. (2014). *Aplikasi model rasch untuk penelitian ilmu-ilmu sosial [Rasch model application for social sciences research]*. Cimahi: Trim Komunikata Publishing House.

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch pada Asesmen Pendidikan*. Cimahi: Trim Komunikata.

Syahrial, S., Ilmah, M., Yahmin, Y., Munzil, M., & Muntholib, M. (2023). Remediation of chemistry teachers' misconceptions about covalent bonding using cognitive conflict interviews: A case study. *Journal of the Serbian Chemical Society*, 88(2), 211–221. Retrieved from https://doi.org/10.2298/JSC220117073S

Yuhanna, W. L., Al Muhdhar, M. H. I., Gofur, A., & Hassan, Z. (2021). Self-reflection assessment in vertebrate zoology (Sravz) using rasch analysis. *Jurnal Pendidikan IPA Indonesia*, 10(1), 35–47. Retrieved from https://doi.org/10.15294/jpii.v10i1.25603

Zhao, C., Zhang, S., Cui, H., Hu, W., & Dai, G. (2023). Middle school students' alternative conceptions about the human blood circulatory system using four-tier multiple-choice tests. *Journal of Biological Education*, 57(1), 51–67. Retrieved from https://doi.org/10.1080/00219266.2021.1877777

Zmnako, S. S. F., & Chalabi, Y. I. (2019). Cross-cultural adaptation, reliability, and validity of the Vertigo symptom scale-short form in the central Kurdish dialect. *Health and Quality of Life Outcomes*, 17(1), 1–12. Retrieved from https://doi.org/10.1186/s12955-019-1168-z

4. Quality of Four-Tier Diagnostic Test - JPFI.pdf

ORIGINALITY REPORT

17%
SIMILARITY INDEX

PRIMARY SOURCES

1	Jumadi Jumadi, Moh Irma Sukarelawan, Heru Kuswanto. "An investigation of item bias in the four-tier diagnostic test using Rasch model", International Journal of Evaluation and Research in Education (IJERE), 2023 Crossref	82 words — 2%
2	Maria Björk, Susanne Knutsson, Elzana Odzakovic, Amanda Hellström et al. " Validation of two brief instruments (the and to measure shared decision-making in patients with restless legs syndrome ", Journal of Sleep Research, 2023 Crossref	42 words — 1%
3	files.eric.ed.gov Internet	30 words — 1%
4	Enoch Teye-Kwadjo. "The Job-Search Self-Efficacy (JSSE) Scale: an Item Response Theory Investigation", International Journal of Applied Positive Psychology, 2021 Crossref	27 words — 1%
5	"Advances in Applications of Rasch Measurement in Science Education", Springer Science and Business Media LLC, 2023 Crossref	23 words — 1%

6	Trikinasih Handayani, Moh Irma Sukarelawan, Mohd Effendi Ewan Mohd Matore, Much Fuad Saifuddin. "Adaptation and evaluation of psychometric properties of HELAM using the Rasch model in the Indonesian context", International Journal of Evaluation and Research in Education (IJERE), 2023 Crossref	22 words — 1%
7	repository.up.ac.za Internet	22 words — 1%
8	"Pacific Rim Objective Measurement Symposium (PROMS) 2015 Conference Proceedings", Springer Science and Business Media LLC, 2016 Crossref	20 words — 1%
9	journal2.uad.ac.id Internet	20 words — 1%
10	Anton J. Slagers, Inge H. F. Reininga, Inge van den Akker-Scheek. "The Dutch language anterior cruciate ligament return to sport after injury scale (ACL-RSI) – validity and reliability", Journal of Sports Sciences, 2016 Crossref	19 words — 1%
11	doktori.bibl.u-szeged.hu Internet	19 words — 1%
12	www.e-iji.net Internet	19 words — 1%
13	kuey.net Internet	18 words — < 1%
14	Antonio P. Gutierrez de Blume, Diana Marcela Montoya Londoño, Virginia Jiménez Rodríguez, Olivia Morán Núñez et al. "Psychometric properties of the	17 words — < 1%

Metacognitive Awareness Inventory (MAI): standardization to an international spanish with 12 countries", Metacognition and Learning, 2024
Crossref

15 www.scribd.com
Internet
17 words — < 1%

16 R N Fardani, C Ertikanto, A Suyatna, U Rosidin. "Practicality and Effectiveness of E-Book Based LCDS to Foster Students' Critical Thinking Skills", Journal of Physics: Conference Series, 2019
Crossref
13 words — < 1%

17 Soeharto Soeharto, Martono Martono, Hairida Hairida, Aigul Akhmetova, Fitria Arifiyanti, Csapó Benő, Charalambous Charalambos. "The metacognitive awareness of reading strategy among pre-service primary teachers and the possibility of rating improvement using Rasch analysis", Studies in Educational Evaluation, 2024
Crossref
13 words — < 1%

18 jppipa.unram.ac.id
Internet
13 words — < 1%

19 journal.stkipsingkawang.ac.id
Internet
12 words — < 1%

20 ejournal.umm.ac.id
Internet
11 words — < 1%

21 ejournal.upsi.edu.my
Internet
11 words — < 1%

22 journal.uinjkt.ac.id
Internet
10 words — < 1%

23 www.iiste.org
Internet
10 words — < 1%

24 Moh. Irma Sukarelawan, Sriyanto Sriyanto, Ariati Dina Puspitasari, Dwi Sulisworo, Ulfiana Nurul Hikmah. "Four-Tier Heat and Temperature Diagnostic Test (4T-HTDT) to Identify Student Misconceptions", JIPFRI (Jurnal Inovasi Pendidikan Fisika dan Riset Ilmiah), 2021
Crossref
9 words — < 1%

25 ejournal.undiksha.ac.id
Internet
9 words — < 1%

26 journal.iainkudus.ac.id
Internet
9 words — < 1%

27 journal.unha.ac.id
Internet
9 words — < 1%

28 journals.vilniustech.lt
Internet
9 words — < 1%

29 jurnal.uns.ac.id
Internet
9 words — < 1%

30 kimika.pfcs.org.ph
Internet
9 words — < 1%

31 vdocument.in
Internet
9 words — < 1%

32 www.science.gov
Internet
9 words — < 1%

33 Achmad Samsudin, Nurul Azizah, Nuzulira Janeusse Fratiwi, Andi Suhandi et al. "Development
8 words — < 1%

of DIGaKiT: identifying students' alternative conceptions by Rasch analysis model", Journal of Education and Learning (EduLearn), 2024
Crossref

34	Francisco J. Moral. "Analysis of soil fertility and its anomalies using an objective model", Journal of Plant Nutrition and Soil Science, 05/03/2012 Crossref	8 words — < 1%
35	Risya Pramana Situmorang, Hadi Suwono, Munzil Munzil, Hendra Susanto, Chun-Yen Chang, Shan-Yu Liu. "Learn biology using digital game-based learning: A systematic literature review", Eurasia Journal of Mathematics, Science and Technology Education, 2024 Crossref	8 words — < 1%
36	Shifi Syarifa Fahmina, Mohammad Masykuri, Dimas Gilang Ramadhani, Sri Yamtinah. "Content validity uses Rasch model on computerized testlet instrument to measure chemical literacy capabilities", AIP Publishing, 2019 Crossref	8 words — < 1%
37	ejournal.unp.ac.id Internet	8 words — < 1%
38	idr.uin-antasari.ac.id Internet	8 words — < 1%
39	vli-journal.org Internet	8 words — < 1%
40	Nurul Miftakhul Janah, Ishafit, Wipsar Sunu Brams Dwandaru. "Simple vertical upward motion experiment using smartphone based phyphox app for physics learning", AIP Publishing, 2024 Crossref	7 words — < 1%

41 Jindong Chang, Yan Li, Hanbing Song, Liming Yong, Lin Luo, Zhanjia Zhang, Naiqing Song. "Assessment of Validity of Children’s Movement Skill Quotient (CMSQ) Based on the Physical Education Classroom Environment", BioMed Research International, 2020
Crossref

6 words — < 1%

42 Muhammad Ramdhan, Moh. Irma Sukarelawan, M. Anas Thohir, Fitria Arifiyanti. "Junior high school student perception of online learning in pandemic Covid-19: Gender, social media ownership, and internet access duration perspective", International Journal of Education and Learning, 2022
Crossref

6 words — < 1%

EXCLUDE QUOTES	OFF	EXCLUDE SOURCES	OFF
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	OFF