International Journal of Evaluation and Research in Education (IJERE)
Vol. 13, No. 6, December 2024, pp. 4279~4286
ISSN: 2252-8822, DOI: 10.11591/ijere.v13i6.29285

Psychometric properties of Indonesian physics motivation questionnaire using Rasch model

Moh Irma Sukarelawan[1], Muhammad Firman Ramadhan[2], Raden Oktova[1], Jumadi[3], Ariati Dina Puspitasari[1], Toni Kus Indratno[1], Zalik Nuryana[4]
[1]Department of Physics Education, Faculty of Teacher Training and Education, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[2]Department of Physics Education, Faculty of Teacher Training and Education, Universitas Muhammadiyah Mataram, Mataram, Indonesia
[3]Department of Physics Education, Faculty of Mathematics and Science, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia
[4]Department of Curriculum and Instruction, School of Education Science, Nanjing Normal University, Jiangsu, China

Article Info

Article history:

Received Nov 17, 2023
Revised Apr 4, 2024
Accepted May 7, 2024

Keywords:

Construct validity
Instrument adaptation
Physics motivation
Psychometric properties
Rasch model

ABSTRACT

The transition to online learning during the pandemic has significantly impacted Indonesian students' physics learning motivation. Recognizing the need for culturally relevant tools to assess this motivation, this study aims to adapt the physics motivation questionnaire (PMQ) to the Indonesian context. This research, which was conducted on 107 secondary and vocational school students, had a balanced gender distribution and used survey methodology. The PMQ, which was adapted from the biology motivation questionnaire (BMQ) through a rigorous forward-backward translation process, was analyzed for its psychometric properties using the Rasch model. This analysis focuses on the content aspects of construct validity, assessing content relevance, representativeness, and technical quality. The results show that the PMQ, using a 4-point Likert scale, effectively addresses these elements, thus confirming its reliability and applicability in the Indonesian educational environment. This study concludes that the culturally adapted PMQ is a valuable instrument for evaluating Indonesian students' motivation to study physics, especially in the context of ongoing changes in the educational landscape.

Corresponding Author:

Moh Irma Sukarelawan
Department of Physics Education, Faculty of Teacher Training and Education, Universitas Ahmad Dahlan Yogyakarta, Indonesia
Email: irma.sukarelawan@pfis.uad.ac.id

1. INTRODUCTION

Developing cognitive and non-cognitive tests is complex, necessitating an extensive literature review and demanding substantial time [1]. This task becomes increasingly challenging as the demand for culturally sensitive instruments rises exponentially in various fields. Instruments designed for one cultural context often require adaptation for others to ensure relevance and efficacy. The growing trend in adapting these instruments across different disciplines and cultural contexts is evidenced by recent studies [2]–[6].

However, adaptation involves more than mere translation; it requires careful consideration of the socio-cultural nuances of the target audience. Established adaptation procedures, such as those proposed by Hambleton [7] and Lee *et al.* [8], offer systematic approaches, yet they also reveal the complexity of this process. The onset of the pandemic has accelerated a shift in educational modes, transitioning abruptly from traditional face-to-face learning to fully online environments. This sudden change has impacted educators and students, altering physical and psychological learning spaces and consequently affecting student motivation [9]. Research indicates that motivated students are more effective in independent learning contexts [10], [11]. Thus, assessing and understanding learning motivation becomes crucial, particularly in these changed circumstances.

The science motivation questionnaire (SMQ) by Glynn *et al.* [12], a tool with proven efficacy in both science and non-science contexts, has undergone significant revisions from its original 30-item, five-dimension format to the more streamlined SMQII with 25 items across five factors [12], [13]. This instrument's versatility is demonstrated through its adaptations in diverse cultural settings, including Greece [14], Turkey [15], Spain [16], the Czech Republic [17], and China [18]. In Indonesia, specific adaptations have been made for dental students by Wardhany *et al.* [19] and for science students by Aini *et al.* [20]. However, while informative, these adaptations have yet to comprehensively address the psychometric properties of SMQII, particularly in the context of the Rasch model and Likert scale functionality.

Furthermore, the adaptation of SMQII in specific academic disciplines like Physics remains limited, unlike its applications in Mathematics [10], Chemistry [14], [15], and Biology [17]. For instance, study by Fiorella *et al.* [10] transformed SMQII into the mathematics motivation questionnaire (MMQ), while Salta and Koulougliotis [14] and Dindar and Geban [15] developed chemistry-focused versions. Janštová and Šorgo [17] streamlined the original SMQ into a more concise 14-item biology motivation questionnaire (BMQ). Drawing from these adaptations, our study focuses on tailoring the BMQ for the field of Physics.

Given the practical advantages of shorter questionnaires, such as higher respondent participation rates [17], we aim to balance brevity with comprehensiveness. While classical test theory, used by Janštová and Šorgo [17], provides some insight, it falls short in reporting detailed psychometric properties, such as item-level reliability and Likert scale functionality. To address these gaps, our study employs the Rasch model to provide a more nuanced understanding of the psychometric qualities of motivation questionnaires in Physics, thus building upon and expanding the existing literature.

2. RESEARCH METHOD

This survey-quantitative research adheres to the guidelines for the Rasch model, which suggests that a sample size between 50 and 250 participants is appropriate for achieving a well-fitting model [21]. To ensure data stability with an accuracy of 0.5 logits at a 95% confidence level, the study recruited 107 students from state senior high and vocational schools, with a gender distribution of 66 males (61.7%) and 41 females (38.3%). We excluded two students (1.8%) with extreme scores to preserve the statistical validity of our findings [22], [23].

The physics motivation questionnaire (PMQ) was meticulously adapted from the BMQ [17], undergoing a rigorous forward-backward translation process by an experienced English lecturer specializing in applied linguistics [6]. The biological terminology was carefully contextualized into the domain of physics, ensuring content relevance for the target demographic. The adapted instrument's readability was then validated through a pilot test with a group of students. The PMQ's motivational scale comprises 14 items across three dimensions-motivation (Mo), self-efficacy (SE), and responsibility (Re)-using a 5-point Likert scale to capture responses.

For data collection, we utilized Google Forms, facilitating an efficient and accessible online survey process [24]. Collaboration with local school teachers ensured a smooth distribution of the questionnaire, and we emphasized to students the confidentiality and academic non-consequential nature of their participation, securing voluntary and unbiased responses. Students typically completed the questionnaire within 15 to 20 minutes. The psychometric quality of the PMQ was scrutinized using the Rasch model, with raw data prepared in Microsoft Excel and analyzed through Winsteps software [25]. This process evaluated the PMQ's construct validity, focusing on content relevance, representativeness, and technical quality [26]. Additionally, we investigated the functional use of the Likert scale, an essential factor in the interpretability of the PMQ results.

3. RESULTS AND DISCUSSION

3.1. Content relevance

Content relevance is critical in understanding the relationship between test items and the measured construct, particularly student motivation in learning physics. To accurately gauge this, the PMQ was adapted from the well-established BMQ, a process meticulously undertaken by Janštová and Šorgo [17]. This adaptation involved tailoring the BMQ's content to reflect the unique aspects of physics, ensuring that the questionnaire resonates with the specific interests and challenges physics students face. The BMQ, known for its robust construction, underwent a rigorous validation process using a factor analysis approach. Out of the original 30 items, 14 were identified as fitting the model criteria, demonstrating the PMQ's strong representation of the construct domain and bolstering its content relevance. This successful adaptation affirms the PMQ’s utility in assessing student motivation in physics and highlights the importance of precise instrument modification in educational research. By capturing the essence of student engagement in physics, the PMQ is a valuable tool in understanding and enhancing the learning experience in this field.

3.2. Representativeness

The strata (H) and separation (G) indices were used to assess the representativeness of the PMQ. These two indexes are displayed in Table 1. In this table, two critical indices, strata and separation, were instrumental in assessing the representativeness of the PMQ. These indices showed promising results, specifically designed to evaluate the instrument's ability to differentiate between varying levels of item difficulty and respondents' abilities. The H and G indices, with respective values of 5.14/4.09 and 3.61/2.82, not only exceeded the established cut-off values [27], [28], but underscored the PMQ's nuanced assessment capabilities. The H index, calculated using H=(4×G+1)/3 [29], further reinforced this point. Regarding reliability, the PMQ demonstrated robustness, with person and item reliability indicating a high degree of consistency in respondents' answers and the quality of the test items. These findings are complemented by a confidence level of 89% and 93% for person and item sizes, respectively, and an impressive Cronbach’s alpha score of 0.92 [30], highlighting excellent interaction quality between items and persons.

Further analysis of the PMQ's content representativeness involved examining item quantity, distribution hierarchy, and gaps in item difficulty [26]. Figure 1 reveals an adequate number of items appropriately distributed across various difficulty levels, it also highlights a notable gap between items Mo3 and Mo4 (0.72 logit). This discrepancy warrants further investigation to understand its impact on the PMQ's overall effectiveness. Nevertheless, the findings collectively affirm the PMQ's strong content representativeness, paving the way for its reliable use in educational contexts to measure student motivation in physics.

Table 1. Summary statistic of PMQ

	Measure		Separation	Reliability	Cronbach’s α
	Mean	SD			
Item	0.00	0.66	3.61	0.93	0.92
Person	2.33	1.60	2.82	0.89	

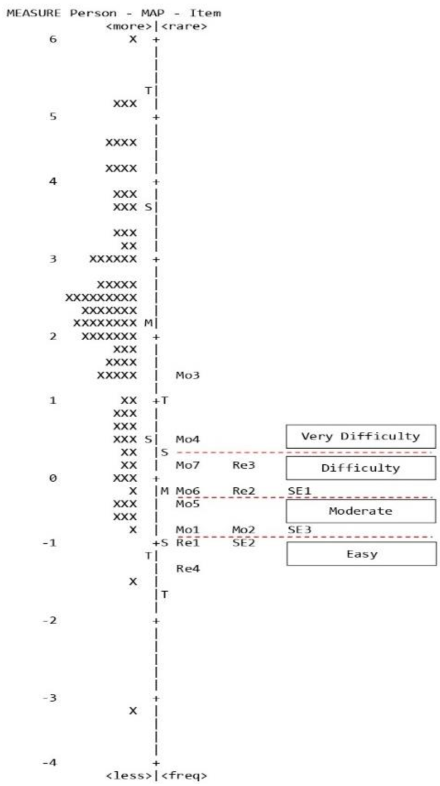

Figure 1. Wright map of student learning motivation

3.3. Technical quality

The fit of empirical data to the Rasch model, particularly through the lens of infit and outfit mean-square (MnSq) values, is a pivotal aspect in assessing the technical quality of the PMQ. While there is no universally accepted standard for statistical fit in Rasch measurements [31], a MnSq value close to 1 is generally considered ideal, with a range of 0.5 to 1.5 often cited as indicative of a good fit [26]. According to data in Table 2, the PMQ generally adheres to this acceptable range, with a notable exception being the Mo3 items, which exhibit slightly elevated MnSq values, above 1.5 but below 2. This deviation suggests that respondents are less responsive to the Mo3 items due to the item's sensitivity or difficulty level. Interestingly, despite this high MnSq value, Mo3's Model S.E. is below 0.30, indicating respondents' strong comprehension of the item. This contrast highlights the nuanced nature of assessing questionnaire validity and suggests that while Mo3 deviates from the ideal fit, its impact on overall understanding may be minimal.

In summary, despite the minor deviation observed in the Mo3 items, the PMQ generally strongly adheres to the criteria for technical quality set by the Rasch model. This suggests that the PMQ is a robust tool capable of effectively measuring student motivation in physics with high technical precision. The overall analysis, therefore, supports the PMQ's utility in educational assessments, with the potential for minor refinements to further enhance its effectiveness.

Table 2. PMQ 14-item model fit

Item	Measure (Logit)	Model S.E (Logit)	Infit		Outfit	
			MnSq	Zstd	MnSq	Zstd
Mo1	-0.42	0.17	1.04	0.33	1.15	1.00
Mo2	-0.48	0.17	1.08	0.65	1.06	0.44
Mo3	1.59	0.15	1.56	3.44	1.62	3.65
Mo4	0.87	0.16	0.65	-2.75	0.64	-2.81
Mo5	-0.15	0.17	0.86	-1.04	0.84	-1.10
Mo6	0.14	0.17	1.21	1.46	1.16	1.11
Mo7	0.44	0.17	1.21	1.43	1.15	1.05
SE1	0.11	0.17	1.06	0.46	1.04	0.34
SE2	-0.54	0.18	1.07	0.55	1.11	0.76
SE3	-0.42	0.17	0.87	-0.96	0.85	-1.00
Re1	-0.70	0.18	0.79	-1.63	0.73	-1.82
Re2	0.14	0.17	0.80	-1.44	0.83	-1.23
Re3	0.44	0.17	0.91	-0.62	0.91	-0.60
Re4	-1.02	0.18	0.74	-2.05	0.69	-1.93

3.4. Structural aspects of construct validity

Unidimensionality, a crucial aspect of test validity, implies that an assessment measures only a single underlying trait [32]–[35]. This concept is particularly significant in evaluating the structural aspects of the PMQ's construct validity [26]. To investigate the unidimensionality of the PMQ, residual principal component analysis (PCAR) is employed [36], [37]. The analysis revealed that the raw variance explained by measures in the PMQ is 51.7%, comfortably surpassing the Rasch recommended threshold of 40% [38]. This high percentage indicates that a substantial portion of the variance in the data can be attributed to the primary latent trait measured by the PMQ, affirming its focus on a single construct. Additionally, the unexplained variance in the first contrast stands at 7.5%, well below the 15% limit, suggesting minimal influence from other dimensions or traits.

These findings underscore the PMQ's unidimensionality, reinforcing its validity as a measure that accurately targets one specific latent trait–student motivation in physics. Such a focused measurement is crucial for ensuring that the PMQ provides meaningful and interpretable results, especially in educational contexts where precise assessment of student motivation is vital. However, it is essential to acknowledge that while PCAR is a robust method for assessing unidimensionality, continuous evaluation and potential refinements to the PMQ should be considered to maintain its relevance and accuracy in evolving educational landscapes.

3.5. Likert rating scale functionality

To complement the assessment of the content aspect of construct validity within the PMQ, the functionality of its 5-point Likert rating scale was scrutinized. This step is critical in confirming that the five rating options are distinct and understandable to the respondents. The analysis, supported by the visual inspection of the response probability curve in Figure 2, indicates a notable deficiency in the respondents' comprehension of the second rank, 'Disagree.' The curve did not achieve a distinct peak for this rank, suggesting that respondents may need to differentiate it clearly from adjacent ranks. This inference is quantitatively bolstered by the Andrich Threshold, which measures the distinctness between adjacent rating categories. The threshold between Likert 2 and Likert 3 stands at a mere 0.38 logits, falling short of the 1.4 logits typically indicative of a clear distinction. A functional curve, as per convention [35], should display a peak at each rating level, signaling a well-understood and effectively utilized scale.

The implications of these findings are significant for the PMQ's utility. The lack of a distinct and functional rating at the 'Disagree' level suggests a potential merging of this category with another to capture the nuances of respondent motivation better. This adjustment could enhance the PMQ’s accuracy in interpreting the degrees of agreement or disagreement expressed by the respondents. Such refinements would increase the reliability of the responses and, consequently, the overall validity of the measurement tool.

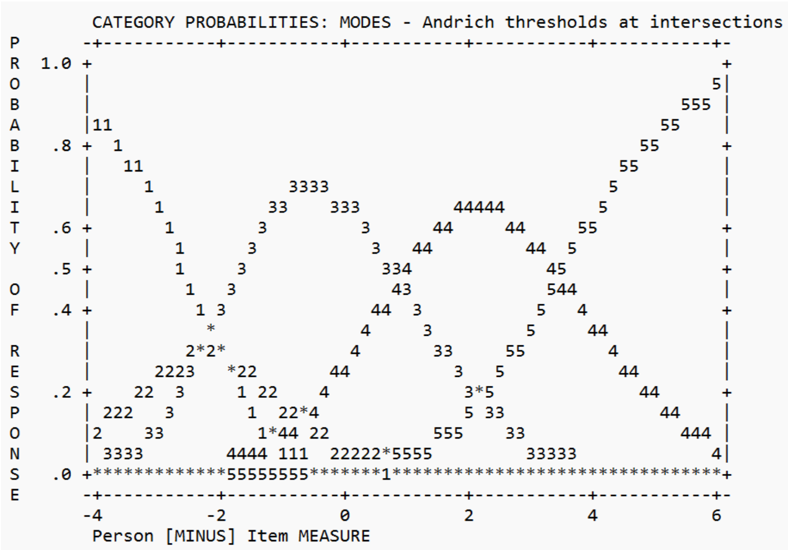

Figure 2. Likert rating scale functionality probability curve in PMQ

4. CONCLUSION

Adapting the PMQ to the Indonesian cultural context has been a substantive endeavor, culminating in successfully representing the construct domain across 14 meticulously adapted items from the BMQ. The psychometric properties of the PMQ are evidenced by the supportive strata (H) and separation (G) indices alongside robust reliability metrics, affirming the test’s representativeness. Further cementing its technical merit, the PMQ’s infit and outfit MnSq values align with the stringent standards of the Rasch model. At the same time, the demonstration of unidimensionality attests to its structural construct validity.

Notably, findings suggest that within the Indonesian cultural milieu, a 4-point Likert scale is more suitable than the original 5-point scale, potentially enhancing the instrument’s sensitivity and response accuracy. This adaptation underscores the PMQ’s applicability and relevance for high school students in Indonesia, offering a valuable tool for assessing and fostering motivation for physics. This research bridges a critical gap by elucidating the quality of a motivational tool in physics education at the high school level, employing the Rasch model for rigorous analysis. However, the scope for further research is ample. Future studies are encouraged to extend the application of the PMQ to broader demographics, including junior high school students and those from private educational institutions, to validate its effectiveness and enrich its cultural adaptability. Such endeavors will not only broaden the instrument’s applicability but also contribute to a more nuanced understanding of student motivation in the diverse educational landscapes of Indonesia.

REFERENCES

[1] T. Handayani, M. I. Sukarelawan, M. E. E. M. Matore, and M. F. 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)*, vol. 12, no. 4, pp. 2362–2373, Dec. 2023, doi: 10.11591/ijere.v12i4.27703.
[2] P. Gouveia, C. Ramos, J. Brito, T. C. Almeida, and J. Cardoso, “The difficulties in emotion regulation scale – short form (DERS-SF): psychometric properties and invariance between genders,” *Psicologia: Reflexao e Critica*, vol. 35, no. 1, p. 11, Dec. 2022, doi: 10.1186/s41155-022-00214-2.
[3] J. P. Uwizihiwe *et al.*, “Cultural adaptation and psychometric evaluation of the Kinyarwanda version of the diabetes-39 (D-39) questionnaire,” *Health and Quality of Life Outcomes*, vol. 20, no. 1, p. 122, Aug. 2022, doi: 10.1186/s12955-022-02034-5.

[4] J. M. A. Peixoto, M. A. V. Gonçalves, M. F. Gaspar, and M. A. V. Matos, "Developing a Valid and Reliable Cross-cultural Measure of Coparenting Conflict between Divorced Parents: The Portuguese Version of the Acrimony Scale," *Journal of Child and Family Studies*, vol. 31, no. 6, pp. 1664–1682, Jun. 2022, doi: 10.1007/s10826-022-02233-0.
[5] M. I. Sukarelawan, Jumadi, H. Kuswanto, Soeharto, and F. N. Hikmah, "Rasch analysis to evaluate the psychometric properties of junior metacognitive awareness inventory in the Indonesian context," *Jurnal Pendidikan IPA Indonesia*, vol. 10, no. 4, pp. 486–495, Dec. 2021, doi: 10.15294/jpii.v10i4.27114.
[6] M. I. Sukarelawan, D. Sulisworo, Jumadi, H. Kuswanto, and S. A. Rofiqah, "Heat and temperature metacognition awareness inventory: A confirmatory factor analysis," *International Journal of Evaluation and Research in Education (IJERE)*, vol. 10, no. 2, pp. 389–395, Jun. 2021, doi: 10.11591/ijere.v10i2.20917.
[7] R. K. Hambleton, "The Next Generation of the ITC Test Translation and Adaptation Guidelines," *European Journal of Psychological Assessment*, vol. 17, no. 3, pp. 164–172, 2001, doi: 10.1027//1015-5759.17.3.164.
[8] W. L. Lee, K. Chinna, K. Lim Abdullah, and I. Zainal Abidin, "The forward-backward and dual-panel translation methods are comparable in producing semantic equivalent versions of a heart quality of life questionnaire," *International Journal of Nursing Practice*, vol. 25, no. 1, pp. 1–9, Feb. 2019, doi: 10.1111/ijn.12715.
[9] D. H. Schunk and M. K. DiBenedetto, "Motivation and social cognitive theory," *Contemporary Educational Psychology*, vol. 60, p. 101832, Jan. 2020, doi: 10.1016/j.cedpsych.2019.101832.
[10] L. Fiorella *et al.*, "Validation of the mathematics motivation questionnaire (MMQ) for secondary school students," *International Journal of STEM Education*, vol. 8, no. 1, p. 52, Dec. 2021, doi: 10.1186/s40594-021-00307-x.
[11] K. A. Renninger and S. Hidi, *The Cambridge handbook of motivation and learning*. Cambridge University Press, 2019.
[12] S. M. Glynn, G. Taasoobshirazi, and P. Brickman, "Science motivation questionnaire: Construct validation with nonscience majors," *Journal of Research in Science Teaching*, vol. 46, no. 2, pp. 127–146, Feb. 2009, doi: 10.1002/tea.20267.
[13] S. M. Glynn, P. Brickman, N. Armstrong, and G. Taasoobshirazi, "Science motivation questionnaire II: Validation with science majors and nonscience majors," *Journal of Research in Science Teaching*, vol. 48, no. 10, pp. 1159–1176, Dec. 2011, doi: 10.1002/tea.20442.
[14] K. Salta and D. Koulougliotis, "Assessing motivation to learn chemistry: Adaptation and validation of Science Motivation Questionnaire II with Greek secondary school students," *Chemistry Education Research and Practice*, vol. 16, no. 2, pp. 237–250, 2015, doi: 10.1039/c4rp00196f.
[15] A. C.-Dindar and Ö. Geban, "Adaptation of the science motivation scale into Turkish and chemistry: Analysis of validity," *Pegem Eğitim ve Öğretim Dergisi*, vol. 5, no. 1, pp. 15–34, Mar. 2015, doi: 10.14527/pegegog.2015.002.
[16] D. Ardura and A. Pérez-Bitrián, "The effect of motivation on the choice of chemistry in secondary schools: Adaptation and validation of the Science Motivation Questionnaire II to Spanish students," *Chemistry Education Research and Practice*, vol. 19, no. 3, pp. 905–918, 2018, doi: 10.1039/c8rp00098k.
[17] V. Janštová and A. Šorgo, "Evaluation, validation and modification of science motivation questionnaire for upper secondary school," *Journal of Baltic Science Education*, vol. 18, no. 5, pp. 748–767, Oct. 2019, doi: 10.33225/jbse/19.18.748.
[18] Z. Dong, M. Li, J. Minstrell, and Y. Cui, "Psychometric properties of Science Motivation Questionnaire II-Chinese version in two waves of longitudinal data," *Psychology in the Schools*, vol. 57, no. 8, pp. 1240–1256, Aug. 2020, doi: 10.1002/pits.22370.
[19] I. I. Wardhany, G. P. Subita, and D. A. Maharani, "Cross-cultural adaptation and psychometric properties of the science motivation questionnaire-II: Indonesian version," *Pesquisa Brasileira em Odontopediatria e Clinica Integrada*, vol. 18, no. 1, pp. 1–10, 2018, doi: 10.4034/PBOCI.2018.181.111.
[20] R. Q. Aini, A. Rachmatullah, F. Roshayanti, S. Shin, J. K. Lee, and M. S. Ha, "Testing validity inferences of science motivation questionnaire (SMQ-II) instrument: Rasch-based analysis with Indonesian secondary students," *Journal of Physics: Conference Series*, vol. 1521, no. 4, p. 042101, Mar. 2020, doi: 10.1088/1742-6596/1521/4/042101.
[21] W. Ling Lee, K. Chinna, and B. Sumintono, "Psychometrics assessment of HeartQoL questionnaire: A Rasch analysis," *European Journal of Preventive Cardiology*, vol. 28, no. 12, pp. E1–E5, Feb. 2021, doi: 10.1177/2047487320902322.
[22] L. Tesio, A. Caronni, D. Kumbhare, and S. Scarano, "Interpreting results from Rasch analysis 1. The 'most likely' measures coming from the model," *Disability and Rehabilitation*, vol. 46, no. 3, pp. 591–603, Jan. 2024, doi: 10.1080/09638288.2023.2169771.
[23] L. Tesio, A. Caronni, A. Simone, D. Kumbhare, and S. Scarano, "Interpreting results from Rasch analysis 2. Advanced model applications and the data-model fit assessment," *Disability and Rehabilitation*, vol. 46, no. 3, pp. 604–617, Feb. 2024, doi: 10.1080/09638288.2023.2169772.
[24] J. Jumadi, M. I. Sukarelawan, and H. Kuswanto, "An investigation of item bias in the four-tier diagnostic test using Rasch model," *International Journal of Evaluation and Research in Education (IJERE)*, vol. 12, no. 2, pp. 622–629, 2023, doi: 10.11591/ijere.v12i2.22845.
[25] J. M. Linacre, "Winsteps® (Version 4.6.1) [Computer Software]." 2021, [Online]. Available: http://www.winsteps.com.
[26] H. T. Ha, "A Rasch-based validation of the Vietnamese version of the Listening Vocabulary Levels Test," *Language Testing in Asia*, vol. 11, no. 1, p. 16, Dec. 2021, doi: 10.1186/s40468-021-00132-7.
[27] J. M. Linacre, *A User's Guide to WINSTEPS® MINISTEP Rash-model computer programs. Program Manual 4.8.0*. 2021.
[28] R. Fisher, Á. Perényi, and N. Birdthistle, "The positive relationship between flipped and blended learning and student engagement, performance and satisfaction," *Active Learning in Higher Education*, vol. 22, no. 2, pp. 97–113, Jul. 2021, doi: 10.1177/1469787418801702.
[29] B. A. Seamon, S. A. Kautz, and C. A. Velozo, "Rasch Analysis of the Activities-Specific Balance Confidence Scale in Individuals Poststroke," *Archives of Rehabilitation Research and Clinical Translation*, vol. 1, no. 3–4, p. 100028, 2019, doi: 10.1016/j.arrct.2019.100028.
[30] F. Firman and S. Rahayu, "Online Learning in the Midst of Covid-19 Pandemic," (in Indonesian), *Indonesian Journal of Educational Science (IJES)*, vol. 2, no. 2, pp. 81–89, Apr. 2020, doi: 10.31605/ijes.v2i2.659.
[31] V. Aryadoust, L. Y. Ng, and H. Sayama, "A comprehensive review of Rasch measurement in language assessment: Recommendations and guidelines for research," *Language Testing*, vol. 38, no. 1, pp. 6–40, Jan. 2021, doi: 10.1177/0265532220927487.
[32] S. Kreiner, "Validity and objectivity: Reflections on the role and nature of Rasch models," *Nordic Psychology*, vol. 59, no. 3, pp. 268–298, Jan. 2007, doi: 10.1027/1901-2276.59.3.268.
[33] C. Hagquist, M. Bruce, and J. P. Gustavsson, "Using the Rasch model in nursing research: An introduction and illustrative example," *International Journal of Nursing Studies*, vol. 46, no. 3, pp. 380–393, Mar. 2009, doi: 10.1016/j.ijnurstu.2008.10.007.

[34] I. S. Simpelaere, G. Van Nuffelen, M. De Bodt, J. Vanderwegen, and T. Hansen, "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*, vol. 15, no. 1, p. 66, Jan. 2017, doi: 10.1186/s12955-017-0639-3.

[35] J. M. Scoulas, B. A. Dunya, and S. L. De Groote, "Validating students' library experience survey using Rasch model," *Library and Information Science Research*, vol. 43, no. 1, p. 101071, Jan. 2021, doi: 10.1016/j.lisr.2021.101071.

[36] R. Shoahosseini and P. Baghaei, "Validation of the Persian translation of the children's test anxiety scale: A multidimensional rasch model analysis," *European Journal of Investigation in Health, Psychology and Education*, vol. 10, no. 1, pp. 59–69, 2020, doi: 10.3390/ejihpe10010006.

[37] S. R.-Madsen, E. E. Wæhrens, B. D.-Samsøe, and K. Amris, "Psychometric properties of the painDETECT questionnaire in rheumatoid arthritis, psoriatic arthritis and spondyloarthritis: Rasch analysis and test-retest reliability," *Health and Quality of Life Outcomes*, vol. 15, no. 1, pp. 1–12, 2017, doi: 10.1186/s12955-017-0681-1.

[38] N. M. Siew and M. S. Abd Rahman, "Assessing the Validity and Reliability of the Future Thinking Test using Rasch Measurement Model," *International Journal of Environmental and Science Education*, vol. 14, no. 4, pp. 139–149, 2019.

BIOGRAPHIES OF AUTHORS

Moh Irma Sukarelawan is an assistant professor at Universitas Ahmad Dahlan (Department of Physics Education, Faculty of Teacher Training and Education), Yogyakarta, Indonesia. He obtained his Doctoral degree in the Department of Educational Science, Graduate School, Universitas Negeri Yogyakarta (UNY) in 2023. His research focuses on physics education, misconception, metacognition, and the Rasch modeling. He can be contacted at email: irma.sukarelawan@pfis.uad.ac.id.

Muhammad Firman Ramadhan is an assistant professor at Universitas Muhammadiyah Mataram (Department of Physics Education, Faculty of Teacher Training and Education), Mataram, Indonesia. He obtained his Doctoral degree in the Department of Educational Science, Graduate School, Universitas Negeri Yogyakarta (UNY) in 2023. His research focuses on physical education, computational physics, physics learning technology.

Raden Oktova is Associate Professor at Universitas Ahmad Dahlan (Department of Physics Education, Faculty of Teacher Training and Education), Yogyakarta, Indonesia. His research focuses on theoretical nuclear physics, general physics, and physics education. He can be contacted at email: r.oktova@pfis.uad.ac.id.

Jumadi is a Professor of Department of Physics Education, Faculty of Mathematics and Science, Universitas Negeri Yogyakarta. Jumadi's research interests lie in the teacher and teacher education, physics education, higher education, 21st century teaching and learning, and e-learning. He can be contacted at email: jumadi@uny.ac.id.

Ariati Dina Puspitasari is an assistant professor at Universitas Ahmad Dahlan (Department of Physics Education, Faculty of Teacher Training and Education), Yogyakarta, Indonesia. Her research focuses on physics education, environmental physics. She can be contacted at ariati.dina@pfis.uad.ac.id.

Toni Kus Indratno is an assistant professor at Universitas Ahmad Dahlan (Department of Physics Education, Faculty of Teacher Training and Education), Yogyakarta, Indonesia. His research focuses on physics education, and physics learning technology. He can be contacted at tonikus@staff.uad.ac.id.

Zalik Nuryana is an assistant professor at Universitas Ahmad Dahlan Indonesia, Islamic Education department. He is currently a Ph.D. student in School of Education Science, Nanjing Normal University China. His areas of interest include curriculum development, teaching method, Islamic education, and online learning. Zalik Nuryana can be contacted at zalik.nuryana@pai.uad.ac.id and zalik@ascee.org; zalik@ieee.org.