Psychometric analysis: Differences in person and item reliability on the adaptation of the Islamic school environment instrument in Indonesia

Mujiansyah[1*], Achmad Ridwan[2], Yuli Rahmawati[3]

[1,2,3]*Jakarta State University, Indonesia.*

Corresponding author: Mujiansyah (*Email: mujiansyah@uin-antasari.ac.id*)

Abstract

This study aims to analyze person and item reliability in the adaptation of an Islamic school environment instrument in Indonesia, as well as evaluate its psychometric quality. The study was conducted at SMAIT Ukhuwah Banjarmasin, involving 80 11th-grade students. The adapted instrument covers dimensions of the physical environment, Islamic foundations, and interaction values in learning. Analysis using the Rasch model shows good overall psychometric properties. Person and item reliability reached 0.97, categorized as "excellent". Infit and Outfit MNSQ values are within an acceptable range, indicating a good fit between data and the model. Unidimensionality analysis yielded a raw variance explained by measures value of 58.5%, indicating the instrument is quite good at measuring a single construct. However, several items were identified as misfit with the Rasch model, indicating the need for revision. High separation values for person (6.13) and item (5.31) demonstrate the instrument's ability to distinguish various trait levels. The results of this study provide a strong foundation for the development and refinement of Islamic school environment measurement instruments in Indonesia.

Keywords: Islamic, Psychometrics, Rasch model, School environment.

Competing Interests: The authors declare that they have no competing interests.

Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

1. Introduction

Islamic education in Indonesia has experienced significant development in recent decades. This development is marked by the emergence of various Islamic schools that strive to integrate Islamic values into the national curriculum [1]. One crucial aspect of ensuring the quality of Islamic education is a supportive school environment. An ideal school environment not only includes physical aspects but also a social and spiritual atmosphere conducive to student development [2]. However,

comprehensive measurement and evaluation of Islamic school environments still face challenges, especially regarding the availability of valid and reliable instruments [3].

Current Islamic school environment measurement instruments often ignore the complexity and uniqueness of local contexts. As a result, measurement outcomes become less representative in depicting actual conditions [4]. Adaptation and validation of instruments that consider the specific dimensions of Islamic school environments in Indonesia become very important to address this issue. These dimensions encompass several important aspects of Islamic education. These aspects include the physical environment, Islamic foundations, teacher-student interaction values, and Islamic values in the classroom learning process [5]. By considering these dimensions, it is expected that measurement instruments can provide a more accurate picture of the Islamic school environment in Indonesia.

The development of this instrument adapts from the Questionnaire on Teacher Interaction (QTI) and What Is Happening In this Class (WIHIC) scales. The QTI, developed by Wubbels et al. [6], focuses on interpersonal teacher behavior and includes dimensions such as leadership, helpful/friendly, understanding, student responsibility/freedom, uncertain, dissatisfied, admonishing, and strict [7]. The WIHIC, created by Fraser et al. [8], assesses the classroom environment through dimensions like student cohesiveness, teacher support, involvement, investigation, task orientation, cooperation, and equity [9]. In adapting these for the Islamic education environment instrument, some dimensions were utilized while others were not included, taking into account the specific context of Islamic schools in Indonesia. This adaptation process aims to create a more culturally and contextually relevant instrument for measuring Islamic school environments.

This research focuses on adapting an Islamic school environment instrument that encompasses various important dimensions. These dimensions include imamah/qiyadah (leadership), ta'awun (mutual assistance), tafahum (understanding), amanah and hikmah (responsibility/wisdom), ukhuwah Islamiyah (Islamic brotherhood), musyarakah (involvement), itqan fi al-'amal (task orientation), and al-'adalah (justice) [10]. The development of such an instrument is expected to provide a deeper understanding of the dynamics of Islamic school environments in Indonesia. By considering these dimensions, the developed instrument can reflect the unique values and practices in Islamic education. Through this instrument adaptation, it is hoped that a more accurate and relevant measurement tool can be obtained to assess the quality of Islamic school environments in Indonesia.

The instrument trial was conducted at the Integrated Islamic High School (SMAIT), Ukhuwah Banjarmasin. This study involved 80 11th-grade students in Mathematics as respondents. The selection of this school and subject was based on SMAIT Ukhuwah's reputation as one of the leading Islamic schools in South Kalimantan. By choosing this school, it is expected to obtain a representative picture of the implementation of Islamic values in the school environment [11]. The psychometric analysis of the results of this trial aims to provide valuable insights into the reliability and validity of the instrument in the context of Islamic education in Indonesia.

Based on this background, the research questions in this study are: (1) How reliable are the person and item measures in the adaptation of the Islamic school environment instrument in Indonesia, particularly at the Integrated Islamic High School (SMAIT) Ukhuwah Banjarmasin? (2) To what extent does the adapted instrument meet good psychometric criteria in the context of Islamic schools in Indonesia? The aim of this research is to analyze the differences in person and item reliability in the adaptation of the Islamic school environment instrument in Indonesia, as well as to evaluate the psychometric quality of the instrument in measuring various dimensions of the Islamic school environment.

The results of this study are expected to make a significant contribution to the development of valid and reliable Islamic school environment measurement instruments in Indonesia. This validated instrument has the potential to become a valuable tool for various parties involved in Islamic education. Educators, policymakers, and researchers can utilize this instrument in efforts to improve the quality of Islamic education in Indonesia [12]. The use of a validated instrument can help identify aspects of the school environment that need improvement. Thus, the results of this study can support data-driven decision-making in the development of Islamic education in Indonesia.

Furthermore, this research is also expected to enrich the literature on psychometric measurement in the context of Islamic education. Currently, the number of studies on psychometric measurement in Islamic education is still limited, both at national and international levels. The contribution of this research can fill the existing knowledge gap in this field. The research results can serve as a reference for other researchers who wish to conduct similar studies in the future. Therefore, this research is not only practically beneficial but also contributes to the development of scientific knowledge in the field of psychometrics in Islamic education.

2. Research Method

This study employs a quantitative approach with a psychometric research design to test the reliability and validity of the adapted Islamic school environment instrument. This approach was chosen for its ability to generate data that can be statistically analyzed to evaluate the psychometric properties of the instrument [13]. The research sample consists of 80 11th-grade students at SMAIT Ukhuwah Banjarmasin in Mathematics class, selected using a *purposive sampling* technique to ensure adequate representation of the target population [14].

The instrument used in this study is an adaptation of various existing school environment measurement scales, modified to include dimensions specific to the Islamic school environment. The instrument adaptation process involved several stages, including translation, back-translation, expert review, and cognitive debriefing to ensure conceptual and linguistic equivalence [15]. The final instrument consists of several aspects, sub-aspects, and indicators measuring dimensions of the physical environment, Islamic foundations, teacher-student interaction values, and Islamic values in the classroom learning process. The instrument data is as shown in the following table:

Table 1.
Islamic-Characteristic General School Instrument.

No	Aspects	Sub-aspects	Indicator/item
1	School physical environment	Learning Facilities and Conditions	6
2	Islamic foundation	Implementation of Islamic Values in Learning Activities	6
3	Islamic values in teacher-student interaction	Imamah/Qiyadah (Leadership)	6
4		Ta'awun (Mutual Assistance)	6
5		Tafahum (Understanding)	6
6		Amanah and Hikmah (Responsibility/Wisdom)	6
7	Islamic values in the classroom learning process	Ukhuwah Islamiyah (Islamic Brotherhood)	6
8		Musyarakah (Involvement)	6
9		Itqan fi al-Amal (Task Orientation)	6
10		Ta'wun (Cooperation)	6
11		Al-'Adalah (Justice)	6
		Total of items	66

Table 2.
Islamic-Characteristic General School Instrument.

No	Sub-aspects	Item	Item total for indicator
1	Learning Facilities and Conditions	The learning spaces at school have ample room for both individual and group work	6
2	Implementation of Islamic Values in Learning Activities	I always feel that religion provides a source of strength in achieving success in learning	6
3	Imamah/Qiyadah (Leadership)	The Islamic teacher (Ustadz/Ustadzah) explains things using mau'izhah hasanah (good teaching/advice)	6
4	Ta'awun (Mutual Assistance)	The Islamic teacher (Ustadz/Ustadzah) demonstrates akhlaq al-karimah (noble character/best behavior) when interacting with students	6
5	Tafahum (Understanding)	The Islamic teacher (Ustadz/Ustadzah) opens up space for mudzakarah (discussion)	6
6	Amanah and Hikmah (Responsibility/Wisdom)	The Islamic teacher (Ustadz/Ustadzah) provides time for tafakkur (reflection) for students during the learning process	6
7	Ukhuwah Islamiyah (Islamic Brotherhood)	I consider all members of this class as ikhwan/akhwat (brothers/sisters)	6
8	Musyarakah (Involvement)	The Islamic teacher (Ustadz/Ustadzah) gives su'al (questions) to me to improve understanding	6
9	Itqan fi al-Amal (Task Orientation)	I complete the wajibat (assignments/duties) to the best of my ability	6
10	Ta'wun (Cooperation)	I apply the principle of shura (consultation/mutual discussion) in completing group projects in class	6
11	Al-'Adalah (Justice)	I receive fair mu'amalah (treatment) along with other students in the class.	6
		Total	66

Data collection in this study was conducted through a self-administered survey given to respondents in controlled sessions within the school environment. This procedure was chosen to ensure a high response rate and minimize potential bias in questionnaire completion [16]. Before starting the questionnaire, respondents were given a comprehensive explanation of the research objectives and asked to provide written consent as a form of willingness to participate in the study. This step was important to meet the ethical aspects of research and ensure that respondents fully understood the implications of their participation. During the questionnaire completion process, researchers were present to answer questions or provide clarification if needed, while maintaining appropriate distance to avoid undesired influence on participants' responses.

Data analysis in this study applied the Item Response Theory (IRT) approach, specifically the Rasch model, which allows for a more in-depth evaluation of the instrument's psychometric properties compared to the Classical Test Theory (CTT)

approach [17]. The Rasch model was chosen for its ability to produce invariant measurements and allow more detailed analysis of item function and individual responses [18]. The advantage of the Rasch model lies in its ability to generate item parameter estimates that are independent of the sample and respondent ability estimates that are independent of the set of items used. Analysis using the Rasch model also allows for the identification of items or respondents that do not fit the model, providing valuable information for instrument improvement. Additionally, the Rasch model facilitates the transformation of ordinal data into interval data, increasing measurement precision and enabling the use of more sophisticated parametric statistical techniques.

The software used for Rasch analysis in this study is Winsteps version 5.8.0, which allows for the calculation of various fit statistics, such as Infit and Outfit Mean Square (MNSQ) and standardized fit statistics (ZSTD) [19]. The analysis includes a comprehensive evaluation of person and item reliability, separation, unidimensionality, and misfit response patterns. The fit criteria used in this analysis refer to standards established in psychometric literature, particularly following the guidelines recommended by Wright [20]. The use of Winsteps software enables the visualization of analysis results through various graphs and tables, facilitating data interpretation. Additionally, Winsteps also provides features to identify problematic items and respondents, allowing researchers to conduct further evaluations of instrument quality.

This research does not rely solely on quantitative analysis but also involves qualitative evaluation of items that show misfit or have unexpected psychometric characteristics. This qualitative evaluation process involves an in-depth review by a panel of experts in the fields of Islamic education and psychometrics to understand the potential causes of item issues. The expert panel is also tasked with proposing item revisions or eliminations if necessary, based on theoretical and practical considerations [21]. This mixed-method approach is expected to provide a more comprehensive understanding of the instrument's psychometric quality. The combination of quantitative and qualitative analyses allows researchers to not only identify problems in the instrument but also understand the context and reasons behind these problems. With this approach, instrument improvements can be made more accurately and effectively, enhancing the overall validity and reliability of the measurement.

3. Results and Discussion

The results of psychometric analysis using the Rasch model indicate that the adapted Islamic-characteristic general school environment instrument has good overall psychometric properties. The mean measure for persons of 2.49 logits, which is higher than the mean measure for items (0.00 logits), suggests that the ability of students in general schools with Islamic characteristics at the high school level is above the average quality of the developed instrument [17]. This finding shows that the instrument is sufficiently sensitive to measure variations in students' perceptions of their Islamic school environment. This instrument's sensitivity allows researchers to identify small differences in student perceptions, which can provide valuable information for the development and evaluation of Islamic education programs. These results also indicate that the instrument has an adequate measurement range to accommodate various levels of student perceptions of their Islamic school environment.

Person and item reliability show excellent results, with a value of 0.97 for both. According to the criteria set by Fisher [22], reliability values above 0.94 are categorized as "excellent". This indicates that the instrument has very high internal consistency and is able to differentiate well between respondents with different trait levels, as well as between items with different difficulty levels [18]. This high internal consistency shows that the items in the instrument work well together to measure the same construct, namely perceptions of the Islamic school environment. The instrument's ability to distinguish between respondents and items with different characteristics demonstrates that this instrument has good measurement precision and can be relied upon to produce accurate data about students' perceptions of their Islamic school environment. As shown in the following table:

Table 3.
Reliability.

Reliability	Description
<0.67	Weak
0.67 – 0.80	Moderate
0.81 - 0.90	Good
0.91 – 0.94	Very Good
>0.94	Excellent

The Infit Mean Square (IMNSQ) and Outfit Mean Square (OMNSQ) values for persons and items in this study are within an acceptable range. Person IMNSQ (1.07) and item IMNSQ (0.98), as well as person OMNSQ (1.28) and item OMNSQ (1.18) are close to the ideal value of 1.00, indicating a good fit between the data and the Rasch model (Linacre, 2021). These values suggest that responses to items in the instrument are fairly consistent and can be well predicted by the model. This fit reinforces the construct validity of the instrument, demonstrating that the items function as expected in measuring students' perceptions of the Islamic school environment. Furthermore, these good fit values also indicate that the instrument is relatively free from statistical noise that could interfere with the interpretation of measurement results. This is supported by the following table:

Table 4.
Infit Mean Square (IMNSQ) and Outfit Mean Square (OMNSQ)Values.

No	Mean-square	Description
1	Value 1.0	Expected ideal value (Fit with model)
2	Less than 1.0	Indicates observations that are too predictable (redundancy, data overfits the model)
3	Value greater than 1.0	Indicates uncertainty (modeled noise, data underfit the model)

Analysis of the Infit and Outfit ZSTD values in this study shows satisfactory results. These values are close to 0, with Infit ZSTD of -0.2 for person and -1.0 for item, and Outfit ZSTD of 0.3 for person and -0.9 for item. According to Wright and Linacre (1994), ZSTD values within the range of -2 to +2 are considered indications of good fit with the Rasch model. These results strengthen the conclusion that the instrument has good measurement quality. ZSTD values close to zero indicate that the response patterns of respondents and item characteristics align with the expectations of the Rasch model, demonstrating consistency in measurement. As shown in the following table:

Table 5.
Infit and Outfit ZSTD Values.

NO	Zstd	Description
1	Value 0.0	The ideal expected value (fit with the model)
2	Less than 0.0	Indicates too easy to predict
3	More than 0.0	Indicates lack of predictability

The separation values obtained for person (6.13) and item (5.31) demonstrate the instrument's ability to distinguish various trait levels among respondents and difficulty levels among items. Higher separation values indicate that the instrument can identify more groups or levels within the sample [22]. In this context, the instrument shows a good ability to differentiate various levels of students' perceptions of their Islamic school environment.

Item analysis involves summing the mean and standard deviation. Infit Mean Square = 0.98 + 0.84 = 1.82. If the logit value per item is greater than 1.82, the item is categorized as not fitting the Rasch model. Based on the item analysis, the items that do not fit the Rasch model are: B11 (4.66), A1 (4.69), A4 (3.13), A3 (2.37), J59 (2.31), and A2 (2.17). The solution for items that do not fit the Rasch model is to update, modify, or remove them. According to Boone et al. [17], misfit items need further evaluation and may require revision or deletion. In this context, a qualitative review of these items is necessary to understand the causes of misfit and to consider improvements or item removal if required.

The unidimensionality of the instrument, measured through raw variance explained by measures, shows good results with a value of 58.5%. According to criteria proposed by Linacre (2021), values between 40% and 59% are categorized as "good." This finding indicates that the instrument is quite effective at measuring the intended single construct, which is the Islamic school environment. This means that the unidimensionality requirement of this data has been met, i.e., good, with raw variance for person at 34.6% (moderate) and for item at 23.8% (moderate). This can be clarified according to the following table:

Table 6.
Unidimensionality Criteria.

Unidimensionality Criteria for Raw Variance	Description
<20%	Weak
21% - 39%	Moderate
40% - 59%	Good
>60%	Very Good

Analysis of the instrument based on student/respondent/person considerations was conducted by summing the mean and standard deviation (SD) of the Infit Mean Square, resulting in a value of 2.25 (1.28 + 0.97). A logit value per person greater than 2.25 indicates that the respondent's or student's data does not fit the Rasch model. Based on the analysis, seven students were identified as not fitting the Rasch model: person 37 (2.32), 52 (2.32), 73 (2.32), 25 (2.45), 40 (2.45), 57 (2.45), and 76 (2.45). Identification of these misfit respondents is important for evaluating data quality and considering potential causes of response inconsistencies. These results indicate the need for further investigation into the response patterns of these seven students to understand factors that may influence their misfit with the Rasch model.

Analysis of respondents' answer patterns in the instrument shows that the lowest score was obtained by student number 8, while the highest score was achieved by student number 27. Answers categorized as similar were identified for students number 27, 42, 59, 78, 43, 60, and 79. A consistent answer trend was observed in the range of questions 5 to 56. These answer patterns provide valuable information about the variation in student responses and consistency in answering a series of specific questions. Further analysis of these answer patterns can help in understanding item characteristics and identifying potential biases or issues in the instrument.

Analysis of item difficulty levels in the instrument resulted in a categorization of questions based on their difficulty. The questions were grouped into easy, medium (in the middle), and difficult categories. These categories help to understand the distribution of difficulty levels in the instrument as a whole. Questions with difficulty levels below average are considered

easy, while those above average are categorized as difficult. This grouping is important for evaluating the instrument's balance in measuring various levels of student ability.

Table 7.
Overview of Item Quality.

Easy Items		Difficult Items			Average Items		Below Average			Above Average		
B8	B9	B11	A1	A3	D22	H47	D19	D23	D24	H46	A5	J59
B10	B7				H48	I51	E25	E26	F32	A2	H45	H44
					J55	K62	F36	G37	H43	A6	D21	F35
							I49	I53	I54	K61	F31	E27
							J58	K64	K65	E29	G40	F33
							K66	C17	E28	F34	K63	J56
							E30	G38	G41	J60		
							C15	C16	C18			
							D20	G42	I50			
							I52	B12	C13			
							C14	G39	J57			

The overall psychometric analysis results indicate that the adaptation of the Islamic school environment instrument has good measurement properties. This instrument demonstrates high reliability and adequate construct validity, as shown by the good fit values in the Rasch model. However, several misfit items and variations in item difficulty levels indicate the need for further improvements to enhance the instrument's quality. These improvements may include revising problematic items or adding new items to strengthen certain dimensions in the instrument. The balance between easy, moderate, and difficult items also needs to be considered to ensure the instrument can accurately measure various levels of students' perceptions of the Islamic school environment.

To validate these findings and explore the instrument's measurement invariance, further research with larger and more diverse samples is needed. This follow-up study is important to test the consistency of the instrument's psychometric properties across various Islamic school contexts in Indonesia. Testing measurement invariance will help ensure that the instrument can be validly used across different student groups and types of Islamic schools. Additionally, further research can explore the relationship between perceptions of the Islamic school environment and other educational variables, such as academic achievement or student character development. Finally, developing score interpretation norms for this instrument will enhance its usefulness in practical contexts within Islamic schools in Indonesia.

4. Conclusion

This research has successfully analyzed the reliability of persons and items in the adaptation of the Islamic school environment instrument in Indonesia, particularly at the Integrated Islamic High School (SMAIT) Ukhuwah Banjarmasin. The analysis results show that the instrument has good overall psychometric properties, with very high person and item reliability (0.97), categorized as "excellent". The Infit and Outfit Mean Square (MNSQ) values for persons and items are within an acceptable range, indicating a good fit between the data and the Rasch model. The unidimensionality analysis of the instrument also shows good results with a Raw variance explained by measures value of 58.5%, indicating that the instrument is quite good at measuring the intended single construct. However, several misfit items and variations in item difficulty levels indicate the need for further improvements to enhance the instrument's quality.

The adapted instrument has met good psychometric criteria in the context of Islamic schools in Indonesia. The high separation values for persons (6.13) and items (5.31) demonstrate the instrument's ability to distinguish between various trait levels in respondents and item difficulty levels. The item difficulty analysis produces a categorization of questions that helps understand the distribution of difficulty levels in the instrument as a whole. Nevertheless, the identification of several misfit items and respondents that do not align with the Rasch model indicates the need for further evaluation and improvement in certain aspects of the instrument. Overall, this instrument shows good potential for use in measuring students' perceptions of the Islamic school environment in Indonesia, with some minor adjustments required.

Based on the research findings, it is recommended to revise or modify the items identified as misfits with the Rasch model, particularly items B11, A1, A4, A3, J59, and A2. This revision process should involve a panel of experts in the fields of Islamic education and psychometrics to ensure that the item improvements maintain content and construct validity. It is advisable to conduct a retest of the revised instrument with a larger and more diverse sample to validate its psychometric properties. Further analysis of the response patterns of students identified as misfits with the Rasch model is necessary to understand the factors that may influence these discrepancies.

The development of score interpretation norms for this instrument is needed to enhance its usefulness in practical contexts within Islamic schools in Indonesia. Further research is suggested to explore the measurement invariance of the instrument across various Islamic school contexts in Indonesia, as well as to investigate the relationship between perceptions of the Islamic school environment and other educational variables, such as academic achievement or student character development. Finally, it is recommended to develop a guide for the use of the instrument for educators and policymakers,

which includes result interpretation and strategies to improve the quality of the Islamic school environment based on the findings from this instrument.

References

[1] D. Afrianty, "Islamic education and youth extremism in Indonesia," *Journal of Policing, Intelligence and Counter Terrorism,* vol. 13, no. 2, pp. 191-206, 2018.

[2] M. A. Lubis, M. M. Yunus, and M. A. Embi, "Systematic review: The measurement of Islamic values in education," *International Journal of Academic Research in Business and Social Sciences,* vol. 9, no. 9, pp. 1125-1137, 2019.

[3] S. Suyatno, D. I. Pambudi, A. Mardati, and E. Nuraini, "The education values of Indonesian teachers: Origin, importance, and Its impact on their teaching," *International Journal of Instruction,* vol. 12, no. 3, pp. 633-650, 2019.

[4] R. Raihani, "Education for multicultural citizens in Indonesia: policies and practices," *Compare: A Journal of Comparative and International Education,* vol. 48, no. 6, pp. 992-1009, 2018.

[5] A. Abdurrahman, A. Saregar, and R. Umam, "The effect of Islamic school culture on students' character: A case study in Indonesia," *Journal of Social Studies Education Research,* vol. 12, no. 1, pp. 149-173, 2021.

[6] T. Wubbels, H. A. Créton, and H. P. Hooymayers, "Discipline problems of beginning teachers, interactional teacher behaviour mapped out," presented at the The American Educational Research Association annual meeting, Chicago, 1985.

[7] T. Wubbels and M. Brekelmans, "Two decades of research on teacher–student relationships in class," *International Journal of Educational Research,* vol. 43, no. 1-2, pp. 6-24, 2005.

[8] B. J. Fraser, C. J. McRobbie, and D. L. Fisher, "Development, validation and use of personal and class forms of a new classroom environment instrument," presented at the The Annual Meeting of the American Educational Research Association, New York, 1996.

[9] J. M. Aldridge, B. J. Fraser, and T.-C. I. Huang, "Investigating classroom environments in Taiwan and Australia with multiple research methods," *The Journal of Educational Research,* vol. 93, no. 1, pp. 48-62, 1999.

[10] S. Arif and H. Maulana, "Islamic school culture and its impact on students' academic achievement: A study in Indonesia," *International Journal of Instruction,* vol. 15, no. 1, pp. 571-586, 2022.

[11] U. Hasanah, R. Sulistiawati, and M. Musfiroh, "The implementation of Islamic values in the learning process: A case study at SMAIT Ukhuwah Banjarmasin," *Journal of Islamic Education,* vol. 8, no. 1, pp. 45-62, 2023.

[12] D. Sulisworo, R. Nasir, and I. Maryani, "Identification of teachers' problems in Indonesia on facing global community," *International Journal of Research Studies in Education,* vol. 6, no. 2, pp. 81-90, 2017.

[13] J. W. Creswell and J. D. Creswell, *Research design: Qualitative, quantitative, and mixed methods approaches*, 5th ed. United States: Sage Publications, 2018.

[14] I. Etikan, S. A. Musa, and R. S. Alkassim, "Comparison of convenience sampling and purposive sampling," *American Journal of Theoretical and Applied Statistics,* vol. 5, no. 1, pp. 1-4, 2016.

[15] D. E. Beaton, C. Bombardier, F. Guillemin, and M. B. Ferraz, "Guidelines for the process of cross-cultural adaptation of self-report measures," *Spine,* vol. 25, no. 24, pp. 3186-3191, 2000.

[16] R. Tourangeau, H. Sun, T. Yan, A. Maitland, G. Rivero, and D. Williams, "Web surveys by smartphones and tablets: Effects on data quality," *Social Science Computer Review,* vol. 36, no. 5, pp. 542-556, 2018.

[17] W. J. Boone, J. R. Staver, and M. S. Yale, *Rasch analysis in the human sciences*. United States: Springer, 2014.

[18] T. G. Bond and C. M. Fox, *Applying the Rasch model: Fundamental measurement in the human sciences*, 3rd ed. United Kingdom: Routledge, 2015.

[19] J. M. Linacre, *Winsteps® Rasch measurement computer program user's guide*. United States: Winsteps.com, 2021.

[20] B. D. Wright, "Reasonable mean-square fit values," *Rasch Measurement Transactions,* vol. 8, no. 3, pp. 370-371, 1994.

[21] R. F. DeVellis, *Scale development: Theory and applications*, 4th ed. United States: Sage Publications, 2017.

[22] W. P. J. Fisher, "Rating scale instrument quality criteria," *Rasch Measurement Transactions,* vol. 21, no. 1, pp. 1095-1097, 2007.