



Rasch Calibration and Differential Item Functioning (DIF) Analysis of the School Well-Being Scale for Students

Eka Hermansyah¹, Charyna Ayu Rizkyanti^{2*}, Lukman Nul Hakim³

^{1,2*,3}Faculty of Education, Universitas Islam Internasional Indonesia.

³Badan Riset dan Inovasi Nasional, Indonesia.

*Corresponding Author. Email: charyna.rizkyanti@uiii.ac.id

Abstract: This study aims to examine the quality of the school well-being scale for high school students. This study used a quantitative approach with Rasch model analysis. The validity, reliability, and differential item functioning (DIF) assessment are all part of the psychometric examination. The 165 high school students in Mataram, West Nusa Tenggara province who took part in this research were 40 males and 125 girls, ages 11 to 18 (mean age = 15.87, SD = 1.39). The findings of the Rasch analysis show the instrument passes the validity test but has weak reliability. The statistical analysis of the Cronbach alpha reliability coefficient of the instrument is 0.43. More detailed results are shown by the Rasch model analysis reliability of person on 0.47 and item reliability of 0.98. This result explains the good quality of the instrument, while the reliability of the instrument needs to be strengthened by controlling the demographic factors. The scale can be used by school management to understand well-being of the students, and further for the improvement of school quality.

Article History

Received: 13-01-2024

Revised: 15-02-2024

Accepted: 27-02-2024

Published: 09-03-2024

Key Words:

School Well-Being
Scale; Konu &
Rimpela; Rasch;
Psychometric.

How to Cite: Hermansyah, E., Rizkyanti, C., & Hakim, L. (2024). Rasch Calibration and Differential Item Functioning (DIF) Analysis of the School Well-Being Scale for Students. *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran dan Pembelajaran*, 10(1), 1-11. doi:<https://doi.org/10.33394/jk.v10i1.10759>



<https://doi.org/10.33394/jk.v10i1.10759>

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



Introduction

The rising incidence of mental health problems among teenagers has sparked global problems, prompting scholars worldwide to take notice. The significance of this issue is apparent from observations that indicate a twofold increase in the majority of young adults having pessimal mental well-being for at least a portion of their day in the 2018 to 2020 timeframe, as compared to the 1993 to 1999 timeframe (Udupa et al., 2023). Udupa et al. (2023) discovered that people classified as Generation Z, indicate greater indicator measures of depressive disorders and also self-injury in their adolescent years. Most of them are still struggling with well-being issues as they grow into adults. Based on data provided by the WHO, the global prevalence of depression is predicted to exceed 300 million individuals (Quinlan & Hone, 2020). Furthermore, forecasts indicate that by 2030, mental illness will surpass all other factors and become the primary cause of disability. Consequently, most contemporary learners are anticipated to experience burnout prior to finishing their schooling (Quinlan & Hone, 2020). Addressing problems associated with "morbidity" among the younger generation, the well-being in the educational context becoming the key topic of discussion, especially in global south countries (Hoferichter et al., 2021).

The presence of multiple aspects in well-being leads to variances in the interpretation of meaning and multiple instruments rising (Aulia et al., 2020; Graham et al., 2017). Measurement in this context must be chosen carefully in accordance with the chosen definition which will determine the dimensions or indicators of the variable (Lammers & Badia, 2005). The instrument quality that is used will determine the quality of the



measurement, considering that what is being measured is a psychometric or latent aspect, precision is important (Bond & Fox, 2015; Lammers & Badia, 2005). The clearer the definition for positioning the variable, the clearer the information about the measurement function.

Methodologically, testing is generally carried out to assess instruments and test hypotheses in quantitative research (Sumintono, 2015). Instrument evaluation is done through the calculation of reliability and validity to see the quality of an instrument (Mohajan, 2017; Sumintono, 2015). Reliable and valid instruments provide quality and trustworthy information while reliability describes how consistently repeated measurements yield convincing information (Sumintono, 2015). The good reliability instrument is not to have the same information from the answer, but the generated answers are still within tolerance limits and are similar (OECD, 2013). On the other hand, validity indicates the accuracy or how well the measuring tool can measure the attributes that should be measured (Sumintono, 2015).

In the vignette approach, instrument testing is conducted to find out if individuals with specific backgrounds understand, interpret, or respond differently to the same survey item (OECD, 2013). The results show that there is a possibility of interpretational differences due to translation variations or cultural differences in responding to items, known as different item functioning (DIF). Therefore, there is a possibility of various instruments being misunderstood when answered by different groups of individuals (OECD, 2013). Considering that culture and individual differences concepts are segregating factors for humans (Buško, 2010).

The framework that is commonly used in the educational context is generally an integrated approach to well-being that is not simply considered from a hedonistic or eudaimonic perspective. Positive emotion, bond, relationship, meaning, and achievement (PERMA) by Seligman (2011) and the school well-being model by Konu & Rimpelä (2002) emerged as new standards that the elements of well-being are a complex concept with subjective, psychological, and social aspects (Kern et al., 2015; Kurniastuti & Azwar, 2014; McLellan & Steward, 2015). Meanwhile, Tian's (2008) school well-being concept is a subjective well-being theory adapted from Diener's (1984) subjective well-being theory. Tian's (2008) school well-being model is still being used today and deals with aspects of school satisfaction, and positive and negative emotions at school (Donat et al., 2016; Renshaw & Chenier, 2019; Tian et al., 2016).

In Indonesia, the available literature study reveals that the prevailing interpretation of school well-being in studies performed is rooted in the framework by Konu and Rimpelä (2002), which draws upon Allardt's conceptual framework for the welfare perspective (Hasanah & Sutopo, 2020; Muhid & Ferdiyanto, 2020). Moreover, the notion of school well-being pertains to the level of pupils' pleasure within the school setting, encompassing their happiness and emotional state while at school (Tian et al., 2014).

As a basis and scale that is widely used in Indonesia, the school well-being instrument is usually carried out on the island of Java (Faizah et al., 2020; Muhid & Ferdiyanto, 2020). Research that tests school well-being instruments mostly uses statistical analysis and uses Cronbach's alpha to analyze validity and reliability (Lathifah et al., 2021). DIF analysis is also used in analyzing psychological well-being in Islamic boarding schools (Khusumadewi & Pramesti, 2023). However, there is very limited analysis of DIF in the context of school well-being in Indonesia, which is based on school well-being by Konu & Rimpelä.

The number of well-being models and instruments used by Konu & Rimpelä (2002) which are widely used in Indonesia need to be balanced with the quality of instruments that accommodate this diversity. DIF analysis also needs to be carried out to ensure that there are



no bias items toward individual differences or cultures (Martinková et al., 2017). As a process of improving the quality of the instrument, this paper fills the shortage of pilot studies in areas with low-quality education in Indonesia.

While developing this model, Konu also conducted empirical confirmation about the factor structure of the model (Konu & Rimpelä, 2002). The data for the study were gathered through the School in Finland, with a total of 40,147 participants on levels 8 and 9. The 43 indicators were obtained from the instrument School Well-being Model through the application of confirmatory factor analysis. The model is applicable for generating well-being profiles for both student groups and the whole school. The school well-being profile can identify specific areas where schools might enhance their efforts to foster the well-being of their students.

Some indicators on Konu's School Well-being Model managed to explain a significant portion of the variance in the overall subjective well-being of students (Konu et al., 2002). Enhancement could be attained by formulating indicators based on the conceptual model. Additional development of the topic might be needed through an examination of the correlations between different categories of school well-being and the potential interconnections among factors within these categories. The test indicates the school's well-being that is conceptually distinct from general subjective well-being.

A comparison was made between the well-being of students based on their gender, school levels, and grades. Primary school children reported superior classroom environments, interpersonal connections, and opportunities for self-fulfillment compared to secondary school learners (Konu & Lintonen, 2006). When examining the relationship between gender and academic performance, the primary discovery was that girls and younger pupils at each educational level reported higher levels of school well-being. However, it should be noted that boys exhibited fewer symptoms compared to girls. The test results indicate that the purpose of the School Well-being Profile is to offer schools an assessment tool for evaluating their well-being.

As a rapidly evolving topic and a crucial part of enhancing the quality of education in Indonesia, there is a need for instruments that have been well-tested for their reliability and do not exhibit bias towards specific demographics (World Bank, 2020). Konu & Rimpelä's (2002) framework is utilized in the majority of research on school well-being conducted in Indonesia. However, the scale of school well-being research is mostly conducted in the Java region (Hasanah & Sutopo, 2020; Khusumadewi & Pramesti, 2023; Lathifah et al., 2021). Constructing a scale in other regions of Indonesia is important to be undertaken amidst the diversity in the country. Therefore, this research chooses to test the school well-being scale in the central region of Indonesia, in the city of Mataram, West Nusa Tenggara Province.

The current framework related to well-being is based on several different models proposed by scholars such as Deci and Ryan (2001), Diener (1984), Keyes (2006), Ryff (1989), and their colleagues because the models they construct describe well-being in general or psychological terms. This study aimed to evaluate the psychometric properties of the School Well-being Scale by Konu & Rimpela instrument by applying the Rasch model. It expects to delve into potential different item function (DIF) across demographic variables, such as gender, school level, parent status, religion, and pocket money. Thus, this current research aim is to examine the quality of the well-being instrument by testing validity, reliability, and using DIF analysis to investigate whether any items are interpreted differently by individuals with specific backgrounds.

Research Method

A quantitative approach with Rasch model analysis was used in this research. Rasch model analysis is a psychometric analysis to assess the instrument capacity (Boone, 2016). The researcher used Google form as the media for distributing the questionnaire and used Rasch model analysis to measure the validity and reliability quality of instruments and DIF analysis to find bias possibilities. DIF analysis will help in finding item bias caused by voter background differences.

This research involved 165 students in Mataram City, Wes Nusa Tenggara province. Student selection was carried out through two methods, namely convenience and snowball sampling. Convenience sampling in this research involves people who are available and comfortable to be involved in research (Johnson & Christensen, 2019). Then, willing participants were also asked to distribute questionnaires to people who were determined to be able to fill out the instrument which was part of snowball sampling (Johnson & Christensen, 2019). The main criteria in this research are teenagers who attend junior or senior high school. It is estimated that they are adolescence, which typically starts between the ages of 11 and 18 (mean age = 15.87, SD = 1.39) (Miller, 2011).

In this pilot study, a well-being scale for schools is generated using a method that includes 26 question items, comprising 16 favorable items and 10 unfavorable items, as demonstrated in Table 1. Apart from that, there is demographic data taken in this research which is listed in Table 2. This study aims to examine the quality of well-being scale by Konu et al. (2002), which includes dimensions of having, loving, being, and health status. These adjustments were previously tested for their validity and reliability by Lathifah et al. (2021). The concept of "having" refers to the overall condition of the school, encompassing both the physical surroundings and the internal environment. In the context of social relationships, loving refers to a social environment that fosters learning and takes into account a range of factors, including relationships between students and teachers, peer interactions, cooperation between the home and school, internal decision-making processes, and the overall organizational climate of the establishment. Self-fulfillment, which refers to the desire for personal growth and satisfaction, can be interpreted in a school setting as the school's effort to offer opportunities for self-fulfillment. Health status refers to an individual's overall state of health. These encompass both physical and mental symptoms. Adjustments to the scale are implemented by the modification of the culture in Indonesia.

A validity test determines the accuracy of the instrument. Validity evaluates the precision of measures, which might take the form of surveys or tests (Sumintono & Widhiarso, 2013). The validity test in this research was conducted by obtaining clearance from experts and selecting items based on the item correlation calculation. The reliability test evaluates the consistency of information in providing similar answers (Sumintono & Widhiarso, 2013).

Table 1. Demographic Data of Participant

Characteristic	Total sample (N = 165)	Percentage (%)
Age		
11	1	0,6
12	6	3,6
13	6	3,6
14	5	3,0
15	32	19,4



16	60	36,4
17	43	26,1
18	12	7,3
Gender		
Male	40	24,2
Female	125	75,8
School Level		
Junior High School	142	86,1
Senior High School	23	13,9
Parent Status		
Complete	140	84,8
No Parent	1	0,6
One Parent	9	5,5
Divorce	15	9,1
Religion		
Islam	142	86,1
Christian	6	3,6
Hindu	17	10,3
Allowance		
under 400000	89	53,9
400000 - 800000	55	33,3
more than 800000	21	12,7

Results and Discussion

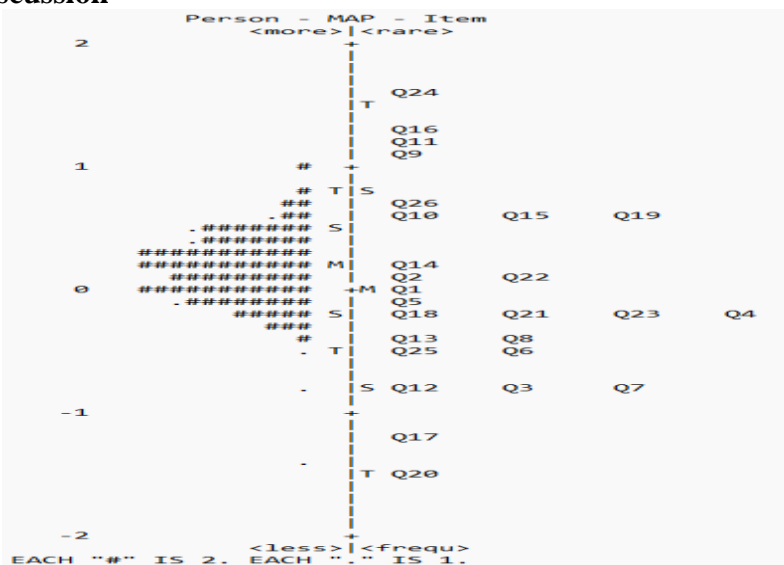


Figure 1. Summary Wright Map of School Well-being

Based on the results of the Wright Map, the distribution of questions is spread well while the students' positions are in the mean area. The distribution of questions is well spread, starting from -1.49 logit on item number 20, and 1.56 logit on item 24 as the most difficult item to

agree on. A good distribution of questions is an even distribution from bottom to top (Sumintono, 2015). Meanwhile, the data distribution shows an approximately normal distribution. It can be seen from the equal data distribution in logit 0 and the mean position (M) as well as the small standard deviation size and focus on the centre of data distribution (Peck et al., 2008). This shows that the school well-being model instrument can be categorized as good and functions optimally in measuring student well-being.

Validity & Reliability

The construct validity analysis using the Rasch model (Figure 1) indicates a raw variance of 34.7%, with the 1st – 5th Unexplained variance percentages at 13.7%, 5.7%, 4.7%, 3.7%, and 3.3%, respectively. Rasch model criteria define values below 20% as less satisfactory, 20-40% as good, 41-60% as very good, and values exceeding 60% as excellent for Raw Variance, while Unexplained Variance values below 15% are acceptable (Widyaningsih & Yusuf, 2018). The results meet the criteria, with good raw variance values (20-40%) and consistently low Unexplained Variance below 15%.

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)				
		-- Empirical --	Modeled	
Total raw variance in observations	=	39.8	100.0%	100.0%
Raw variance explained by measures	=	13.8	34.7%	34.7%
Raw variance explained by persons	=	1.4	3.6%	3.6%
Raw Variance explained by items	=	12.4	31.1%	31.1%
Raw unexplained variance (total)	=	26.0	65.3%	100.0%
Unexplnd variance in 1st contrast	=	5.4	13.7%	20.9%
Unexplnd variance in 2nd contrast	=	2.3	5.7%	8.8%
Unexplnd variance in 3rd contrast	=	1.9	4.7%	7.2%
Unexplnd variance in 4th contrast	=	1.5	3.7%	5.7%
Unexplnd variance in 5th contrast	=	1.3	3.3%	5.0%

Figure 2. Standardized Residual Variance for Validity

As a feature of the Rasch model, the reliability test analysis examines items and individuals. Person Reliability and Item Reliability are categorized as Excellent (>0.94), Special (0.91-0.94), Good (0.81-0.90), Adequate (0.67-0.80), and Weak (<0.67) (Sumintono & Widhiarso, 2015). Figure 2 reveals findings showing a classical Cronbach's alpha value of .43, indicating weak reliability. However, Rasch model analysis shows a higher Person reliability of .47 (weak reliability) and an item reliability of .98 (special reliability), emphasizing the instrument's overall reliability. The weak reliability is attributed to inconsistent student responses, possibly influenced by the online and unsupervised nature of the research. Despite this, the instrument exhibits high-quality items validated through pre-testing.

SUMMARY OF 165 MEASURED Person								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIIT		OUTFIIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	68.7	26.0	.16	.23	1.01	-.2	1.01	-.2
S.D.	6.0	.0	.31	.01	.50	2.0	.52	1.9
MAX.	83.0	26.0	.96	.28	2.98	5.4	3.28	5.7
MIN.	41.0	26.0	-1.39	.22	.15	-5.7	.18	-5.1
REAL RMSE	.25	TRUE SD	.19	SEPARATION	.75	Person	RELIABILITY	.36
MODEL RMSE	.23	TRUE SD	.21	SEPARATION	.94	Person	RELIABILITY	.47
S.E. OF Person MEAN = .02								
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .43								
SUMMARY OF 26 MEASURED Item								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIIT		OUTFIIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	436.0	165.0	.00	.09	1.00	-.1	1.01	.0
S.D.	95.6	.0	.75	.01	.18	1.9	.19	2.0
MAX.	595.0	165.0	1.56	.13	1.38	3.9	1.43	4.0
MIN.	245.0	165.0	-1.49	.08	.68	-4.0	.68	-4.0
REAL RMSE	.10	TRUE SD	.75	SEPARATION	7.73	Item	RELIABILITY	.98
MODEL RMSE	.09	TRUE SD	.75	SEPARATION	8.02	Item	RELIABILITY	.98
S.E. OF Item MEAN = .15								

Figure 3. Rasch Model Analysis Summary

GUTTMAN		SCALOGRAM OF RESPONSES:									
Person	Item	21	12	2	21	11	12	11	2		
	07237653883415122405969164										
75	+33333333333333333333333333333333										
93	+33333333333333333333333333333333										
106	+44444444414444444441111111111111										
116	+44444444424442424322111111111111										
15	+44444444444444141441111111111111										
140	+44141144444414144144141111111111										
155	+44444444144144441441111111111111										
58	+4323222222222222232322222223										
115	+14144414111111111111111111111111										
	07237653883415122405969164										
	211 21 12 2 211112 112										
	07237653883415122405969164										

Figure 4. Scalogram of School Well-being Scale

The scalogram analysis (figure 4) revealed deviations in participant responses, with individuals like 75 and 93 consistently using a score of "3" for various questions, irrespective of difficulty levels. Michalos (2014) suggests that responses on a well-being scalogram should ideally vary in difficulty levels. However, participants 106, 116, 15, 140, and 155 exhibited poor results, consistently choosing extreme points "4" or "1" rather than the expected "3" or "2." Participant 115 tended to provide arbitrary responses, even assigning the easiest questions a score of "1" instead of the expected "4." In contrast, participant 58 consistently used points "2" and "3" across various difficulty levels.

DIF analysis

DIF analysis will verify the low person reliability value caused by students' errors in answering instrument items randomly or there are interpretation errors in certain groups in each question item. It will be carried out on demographic data on gender, school level, parent status, and religion. Probability values that show below 0.05 or 5% are a concern so as not to harm certain groups (Sumintono, 2015).

In the gender category, there is DIF in item number 2 "Over the past six months, I have experienced feeling tired or weak", item number 10 "Over the past six months, I have experienced headaches", item number 15 "Over the past six months, I have experienced neck or shoulder pain", item number 21 "I have felt healthy in the past six months", and item number 26 "Over the past six months, I have experienced stomach pain". The interesting point is that it is all related to the health status dimension. Women tend to easily agree when the question is favorable, and it is more difficult to agree when the question is unfavorable compared to men. However, health status in item number 3 "My hearing is in good condition, and I am healthy in participating in school activities" does not show bias in the answers of students, whether male or female.

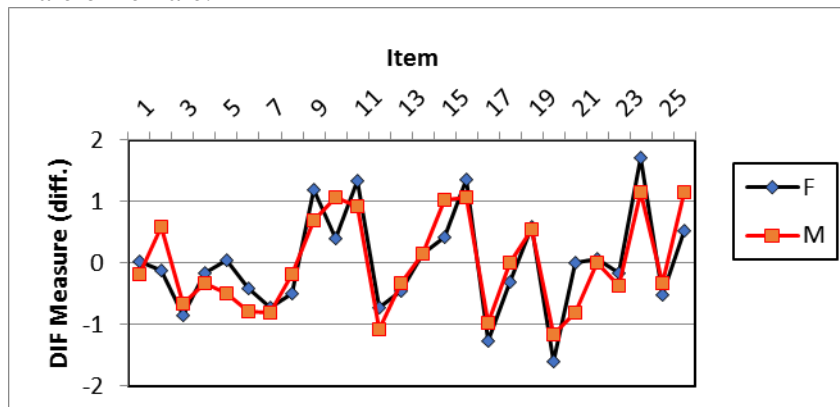


Figure 5. DIF Analysis Based on Gender

The health status point could represent a cause of stress that must be examined in order to evaluate whether the factors in the health status dimension are biased toward women or whether the student's health status is similar. According to Klärner et al. (2022), throughout puberty, girls tend to have chronic and psychosomatic pain that is easier to feel, whereas boys are likely to develop deadly illnesses. This finding also occurred in the English language instrument testing by Konu & Lintonen (2006) which shows that there are more symptoms in adolescent girls than boys.

Meanwhile, at the school level, which is divided into junior & senior high school (labeled j for junior high school & s for senior high school on Figure 6), on the dimension of having item no 8 "The school environment and my classroom are clean and neat" junior high school students find it more difficult to agree than senior high school students. Item number 8 has a position below the student average, it should be easy to agree on this item. On item number 22 "Bullying has never occurred at my school" junior high school students find it more difficult to agree than senior high school students. There are indications that there is a difference in the level of inclusiveness between junior high school levels in Mataram city compared to senior high school.

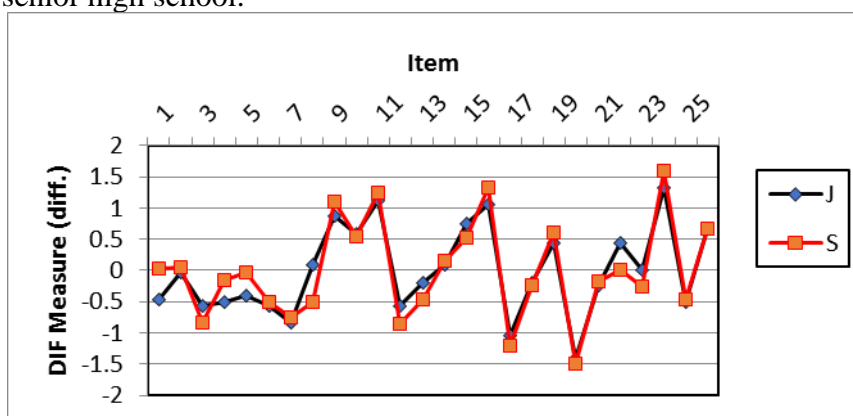


Figure 6. DIF analysis based on School Level

Based on pocket money which is divided into low (<400,000), medium (400,000 – 800,000), and high (>800,000). It was shown that there was no significant DIF in the schoolwell-being instrument based on the amount of student pocket money. Likewise, the parent status category which was divided into complete, no parent, single parent, and divorce showed that there was no significant DIF. This is also followed by the religion category consisting of Islam, Christianity and Hinduism, where each item shows a point above 0.05 as an indicator of no DIF or potential bias in the item towards a particular religion.

Conclusion

In conclusion, the rasch calibration and DIF analysis on the school well-being Scale in Mataram have provided insights into the reliability, validity, and DIF within the instrument. The reliability analysis, considering both person and item reliability, revealed a weak overall instrument reliability, primarily attributed to inconsistencies in student responses, possibly influenced by the unsupervised online administration of the questionnaire. However, the high item reliability indicated the robustness and consistency of the measurement across the items, emphasizing the quality of the instrument itself.

The DIF analysis delved into potential biases across demographic variables, revealing notable gender and school-level-related differences. In terms of gender, specific items related to health status showed bias, with females more likely to agree with favorable statements and less likely with unfavorable ones compared to males. Additionally, differences between



junior and senior high school students were observed in items related to the cleanliness of the school environment and occurrences of bullying, suggesting variations in perceptions between these school levels in Mataram. However, encouragingly, no significant DIF was found based on the variables of pocket money, parent status, and religion, indicating that the instrument performs consistently across these factors. As the study contributes to the ongoing dialogue on measurement in well-being research, it emphasizes the need for careful consideration of demographic factors when interpreting and utilizing well-being assessment tools.

Recommendation

The results showed that the school well-being scale, developed by Konu and Rimpela, is valid, meaning that it measures the well-being of the study participants. The administration of schools and universities can use the scale to gauge how well their students are doing. This is crucial for preventative measures and for assisting the student in adjusting to the school environment and achieving academic success. The Ministry of Education, Culture, Research, and Technology can be urged to give student well-being more consideration by policy makers, such as members of parliament from the tenth commission. Over time, raising student well-being will raise the standard of education in Indonesia as a whole.

References

- Aulia, F., Hastjarjo, T. D., Setiyawati, D., & Patria, B. (2020). Student Well-being: A Systematic Literature Review. *Buletin Psikologi*, 28(1), 1. <https://doi.org/10.22146/buletinpsikologi.42979>
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (Third edition). Routledge, Taylor and Francis Group.
- Boone, W. J. (2016). Rasch Analysis for Instrument Development: Why, When, and How? *CBE—Life Sciences Education*, 15(4), rm4. <https://doi.org/10.1187/cbe.16-04-0148>
- Buško, V. (2010). Measuring individual differences in psychological attributes: A psychometric view of contextual effects. *Review of Psychology*, 17(1).
- Donat, M., Peter, F., Dalbert, C., & Kamble, S. V. (2016). The Meaning of Students' Personal Belief in a Just World for Positive and Negative Aspects of School-Specific Well-Being. *Social Justice Research*, 29(1), 73–102. <https://doi.org/10.1007/s11211-015-0247-5>
- Faizah, F., Rahma, U., Dara, Y. P., & Gunawan, C. L. (2020). School Well-Being Siswa Sekolah Dasar dan Siswa Sekolah Menengah Pertama Pengguna Sistem Full-Day School di Indonesia. *Jurnal Kajian Bimbingan dan Konseling*, 5(1), 34–41. <https://doi.org/10.17977/um001v5i12020p034>
- Graham, A., Powell, M. A., Thomas, N., & Anderson, D. (2017). Reframing 'well-being' in schools: The potential of recognition. *Cambridge Journal of Education*, 47(4), 439–455. <https://doi.org/10.1080/0305764X.2016.1192104>
- Hasanah, M., & Sutopo, S. (2020). Pengaruh School Well-Being terhadap Motivasi Belajar Siswa di Madrasah Aliyah Ma'arif 7 Sunan Drajat Lamongan. *Ummul Qura: Jurnal Institut Pesantren Sunan Drajat (INSUD) Lamongan*, 15(2), 114–123. <https://doi.org/10.55352/uq.v15i2.163>
- Hoferichter, F., Kulakow, S., & Hufenbach, M. C. (2021). Support From Parents, Peers, and Teachers Is Differently Associated With Middle School Students' Well-Being. *Frontiers in Psychology*, 12, 758226. <https://doi.org/10.3389/fpsyg.2021.758226>



- Johnson, R. B., & Christensen, L. B. (2019). *Educational Research: Quantitative, Qualitative, and Mixed Approaches*. SAGE Publications, Inc.
- Kern, M. L., Waters, L. E., Adler, A., & White, M. A. (2015). A multidimensional approach to measuring well-being in students: Application of the PERMA framework. *The Journal of Positive Psychology*, 10(3), 262–271. <https://doi.org/10.1080/17439760.2014.936962>
- Khusumadewi, A., & Pramesti, M. (2023). Development of Psychological Well-Being Scale for Students in Islamic Boarding Schools. In A. Mustofa, I. Widiyanah, B. K. Prahani, I. A. T. Rahayu, Moh. Mudzakkir, & C. D. M. Putri (Eds.), *Proceedings of the International Joint Conference on Arts and Humanities 2023 (IJCAH 2023)* (Vol. 785, pp. 1098–1104). Atlantis Press SARL. https://doi.org/10.2991/978-2-38476-152-4_111
- Klärner, A., Gamper, M., Keim-Klärner, S., Moor, I., Von Der Lippe, H., & Vonneilich, N. (Eds.). (2022). *Social Networks and Health Inequalities: A New Perspective for Research*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-97722-1>
- Konu, A., & Lintonen, T. P. (2006). School well-being in Grades 4-12. *Health Education Research*, 21(5), 633–642. <https://doi.org/10.1093/her/cyl032>
- Konu, A., Lintonen, T. P., & Rimpela, M. (2002). Factors associated with schoolchildren's general subjective well-being. *Health Education Research*, 17(2), 155–165. <https://doi.org/10.1093/her/17.2.155>
- Konu, A., & Rimpelä, M. (2002). Factor structure of the School Well-being Model. *Health Education Research*, 17(6), 732–742. <https://doi.org/10.1093/her/17.6.732>
- Kurniastuti, I., & Azwar, S. (2014). Construction of Student Well-being Scale for 4-6th Graders. *Jurnal Psikologi*, 41(1), 1. <https://doi.org/10.22146/jpsi.6954>
- Lammers, W. J., & Badia, P. (2005). *Fundamentals of behavioral research* (1st ed). Thomson/Wadsworth.
- Lathifah, N., Adi, A. H., & Na'imah, T. (2021). An Overview of School Well-Being Scale at Students of Muhammadiyah Vocational High School: A Pilot Study. *Proceedings Series on Social Sciences & Humanities*, 2, 97–104. <https://doi.org/10.30595/pssh.v2i.111>
- Martinková, P., Drabinová, A., Liaw, Y.-L., Sanders, E. A., McFarland, J. L., & Price, R. M. (2017). Checking Equity: Why Differential Item Functioning Analysis Should Be a Routine Part of Developing Conceptual Assessments. *CBE—Life Sciences Education*, 16(2), rm2. <https://doi.org/10.1187/cbe.16-10-0307>
- McLellan, R., & Steward, S. (2015). Measuring children and young people's wellbeing in the school context. *Cambridge Journal of Education*, 45(3), 307–332. <https://doi.org/10.1080/0305764X.2014.889659>
- Michalos, A. C. (Ed.). (2014). *Encyclopedia of Quality of Life and Well-Being Research*. Springer Netherlands. <https://doi.org/10.1007/978-94-007-0753-5>
- Miller, P. H. (2011). *Theories of developmental psychology* (5th ed). Worth Publishers.
- Mohajan, H. K. (2017). Two Criteria for Good Measurements in Research: Validity And Reliability. *Annals of Spiru Haret University. Economic Series*, 17(4), 59–82. <https://doi.org/10.26458/1746>
- Muhid, A., & Ferdiyanto, F. (2020). Stres akademik pada siswa: Menguji peranan iklim kelas dan school well-being. *Persona: Jurnal Psikologi Indonesia*, 9(1), 140–156. <https://doi.org/10.30996/persona.v9i1.3523>
- OECD (Ed.). (2013). *OECD guidelines on measuring subjective well-being*.



- Peck, R., Olsen, C., & Devore, J. L. (2008). *Introduction to statistics and data analysis* (3rd ed). Thomson Brooks/Cole.
- Quinlan, D. M., & Hone, L. C. (2020). *The Educators' Guide to Whole-school Wellbeing: A Practical Guide to Getting Started, Best-practice Process and Effective Implementation* (1st ed.). Routledge. <https://doi.org/10.4324/9780429280696>
- Renshaw, T. L., & Chenier, J. S. (2019). Screening for Student Subjective Well-Being: An Analog Evaluation of Broad and Targeted Models. *Journal of Psychoeducational Assessment*, 37(6), 803–809. <https://doi.org/10.1177/0734282918795797>
- Sumintono, B. (2015). *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan* (W. Widhiarso, Ed.). Penerbit Trim Komunikata.
- Sumintono, B., & Widhiarso, W. (2013). *Aplikasi model rasch: Untuk penelitian ilmu-ilmu sosial*. Trim Komunikata Publishing House.
- Tian, L., Han, M., & Huebner, E. S. (2014). Preliminary development of the Adolescent Students' Basic Psychological Needs at School Scale. *Journal of Adolescence*, 37(3), 257–267. <https://doi.org/10.1016/j.adolescence.2014.01.005>
- Tian, L., Tian, Q., & Huebner, E. S. (2016). School-Related Social Support and Adolescents' School-Related Subjective Well-Being: The Mediating Role of Basic Psychological Needs Satisfaction at School. *Social Indicators Research*, 128(1), 105–129. <https://doi.org/10.1007/s11205-015-1021-7>
- Udupa, N. S., Twenge, J. M., McAllister, C., & Joiner, T. E. (2023). Increases in poor mental health, mental distress, and depression symptoms among U.S. adults, 1993–2020. *Journal of Mood and Anxiety Disorders*, 2, 100013. <https://doi.org/10.1016/j.xjmad.2023.100013>
- World Bank. (2020). *The Promise of Education in Indonesia*. World Bank.