

Rasch Model Analysis of Indonesian Spiritual Well-Being

R. Rostiana^{1*} Zamralita Zamralita¹ Erik Wijaya¹

¹Faculty of Psychology, Universitas Tarumanagara, Jakarta, Indonesia

*Corresponding author. E-mail: rostiana@fpsi.untar.ac.id

ABSTRACT

The existing measurement of spiritual well-being are mostly based on a western perspective. Consequently, its not sufficient to describe the spiritual well-being of easterners, including Indonesia. Considering this fact, this study aims to develop a scale of spiritual well-being based on the perspective of Indonesians. In order to obtain psychometric properties that meet the requirements, validity and reliability tests are conducted based on the Rasch model. Data were collected online through a monkey survey application which contains a scale of Indonesia's spiritual well being containing 35 items. This scale was sent to schools, universities and companies. The collected data (N = 350) were processed to determine the validity and reliability of the scale. The results showed that the scale validity requirements can be met through the unidimensionality test, as well as the infit and outfit score of MNSQ which show that in general the scale of Indonesia's spiritual well being meets the requirements. Based on the item fit test, there were 4 items that must be discarded so that the number of items remaining is 31. Reliability requirements were also met through the Cronbach alpha score, item and person reliability score. However, the Rasch model of the Indonesian Spiritual well being showed that the rating scale still requires further study related to the ambiguity of ratings on this scale

Keywords: spiritual well-being, Rasch model, validity, reliability

1. INTRODUCTION

The hedonic and eudaimonic techniques are commonly used to study the concept of well being [1]. According to [2], the hedonic approach is based of life satisfaction and happiness, or pleasurable events [3]. On the other hand, Aristotle's eudaimonic approach emphasizes empowering potential [4]. Along with these two approaches, some academics learned well being from a spiritual perspective [5], however this perspective is currently under-researched compared to hedonic or eudaimonic perspectives.

Has played a significant role in people's lives in Indonesia, and it is not surprised that the spiritual element also appeared in the researchers' initial study of the meaning of well being. This study's findings indicated some spiritual responses such as 'Closeness to God'; 'Patience' (sabar); 'Gratitude'; and 'Sincerity' (ikhlas). Those responses are not covered either in hedonic or in eudaimonic approached. According to [6], closeness to God and meaning of life are the core of spiritual well being that is theologically oriented (based on religion). As a result, those findings should be followed up with more in-depth research.

Currently, there are still limited available instruments to measure spiritual well being, among others developed by [7], [8], [9], and [10]. They are all arranged in western perspective, and none of them included an element such as gratitude, patience, and sincerity that are frequently

expressed and demonstrated in everyday life of Indonesians. So, it is essential to make an Indonesian version of the spiritual well being scale, based on the study's first findings.

Based on the result of focus group discussion (FGD) attended by 98 people representing diverse demographics (religion, education, age, and gender), well being of Indonesians are contented for a variety of reasons, including their belief in God, their application of patience, gratitude, and sincerity in life, and their meaningful lives. These findings served as the foundation to compose item scale and produced 35 items whose psychometric properties will be examined using Rasch modeling. The purpose of this study is to describe the scale's validity, reliability, and unidimensionality, as well as to conduct a rating scale analysis of the spiritual pleasure scale.

2. BACKGROUND

2.1. The meaning of spiritual well-being

Spirituality, as [11] revealed, is a multidimensional variable that may be interpreted in a variety of ways. For example, [12] emphasized the degree of contentment with the spiritual sphere. According to [5], spirituality is an extraordinary dimension of everyday human life. Spirituality

pervades every significant moment of life, from birth to death, both in times of pleasure and crisis. However, strangely, the significance of spirituality is frequently recognized only when this feature is absent. Loneliness, suffering, and an inability to accept unfavorable events are all instances in which spirituality is required. If it is discovered, the suffering may typically be transferred.

Fisher established the following characteristics of spirituality: 1) It is an inherent characteristic. Spirituality develops from the moment an individual enters the world and improves in quality due to the effect of a supportive environment. 2) Emotional dimension. Individual attitudes become no longer neutral or free to evaluate when spiritual experiences affect the heart; they will always relate to the depth of emotions they feel, such as serenity, tranquillity, kindness, and wisdom (positive emotions) while judging things. According to [11], the feature of connection or transcendence is the primary reference (characteristic) for comprehending spirituality in its entirety. Paloutzian defined spirituality as having two dimensions: (1) a transcendental dimension encompassing human interactions with divine forces and (2) a meaningful life. Similarly, according to [13], spirituality encompasses an individual's relationship with himself, others, the universe, and a transcendental relationship with God. Spirituality is dynamic and is required to positively interpret life experiences to attain the desired quality of life. In spirituality, the connection has two orientations: a vertical (transcendental) direction and a horizontal (individual or social) orientation. Spiritual well-being is distinguished from subjective and psychological well-being, which are only horizontal in direction.

2.2. Classical test theory vs. item response theory

There are two critical issues in measuring psychological attribution [14]. To begin with, psychological assessment must adhere to psychometric standards, specifically the instrument's validity and reliability. Second, statistical scores have consequences for both notions. The validity and reliability of measuring devices are critical factors in determining their quality [15].

There are two measurement theories: Classical True-Score Theory and Modern/Latent-Trait Theory (IRT) [16]. Modern Test Theory appears to address several limitations of Classical Test Theory; In Classical Test Theory, parameters refer to the properties of the items whose computations are contingent on the sample group utilized. In this regard, the degree of difficulty and weighting of items are also determined by the characteristics of the group being tested. Additionally, in traditional test theory, the reliability score is conditional on the sample's condition, non-linear, and constrained to existing scores' range [17]. IRT is formulated to answer the limitations of the Classical Test Theory, such as a high dependence on sample characteristics both from the point of view of the statement items and the characteristics of the sample [18]. Rasch modeling implements Item Response Theory (IRT),

developed by Georg Rasch, who argued that the opportunity to solve a problem depends on the ratio of the individual's ability to the level of difficulty of the question [17]. The Rasch probability principle can be interpreted as follows: Individuals with higher abilities will have a greater chance of answering questions correctly and completing more complicated tasks. Furthermore, Rasch developed a measurement model based on ratio data to determine the relationship between person ability and item difficulty using the logarithmic function to produce measurements with equal intervals. The result is a new unit called logit (log odds unit), which shows a person's ability and item difficulty so that from the logit value obtained, it can be concluded that the individual's level of success in working on questions is highly dependent on the level of ability and level of difficulty of the questions [17]. The implementation of Rasch modeling in social research can contribute to increasing the validity and reliability of measuring instruments.

3. METHOD

3.1. Participant

The participants were 350 people, with the characteristics as follow: The participants are dominated by female (53.4%), with age average 26.98 years; they are not married yet (99.4%) and the education is dominated by high school degree (89.4%).

3.2. Measurement and Data Collecting

The Indonesian Spiritual Well-being scale (ISW) was employed as the scale containing 35 items. The scale assessed the transcendental dimension: an individual's relationship with God or an external resource frequently relied upon for survival. Additionally, the ISW scale examined the spiritual virtues that constitute the foundation for action, notably patience, gratitude, and sincerity. Examples of items included "I believe that my life is sufficient; there is no reason to be jealous of others"; ISW also assessed aspects of life's meaning or purpose. "I have a clear purpose in life" is an example of item. The data were collected with utilizing the Monkey Survey Computer Application. The scale was delivered to many universities, companies and schools in Java. This scale contained informed consent, personal information, and the ISW scale. The analysis yielded the respondent's accuracy index, item accuracy index, respondent-item accuracy map, unidimensionality, and Cronbach's alpha reliability ratings

4. RESULTS

The analysis results that were displayed based on the Rasch model include summary statistics, level of conformity of statement items, unidimensionality, and rating scales. The

data processed were 35 statement items and 350 respondents. The results of data processing and discussion are described below.

4.1. Summary of statistics

Based on the data processing results using the winsteps program, the data analyzed were 34965 with a Chi-squared value of 78132.11 with degrees of freedom (df) 32919 (p = 0.00), indicating that the overall measurement was good and very good. There is one participant with extreme scores or is considered an outlier.

- *Mean person measured* = +1.21 logit (>0 logit) indicates that participants tend to agree more with the measurement items.
- Cronbach's alpha value (KR-20) of 0.91 (> 0.8) indicates that the interaction between the person (participant) and the statement items as a whole is very good.
- The value of person reliability is 0.88, and item reliability is 1.00, indicating that the consistency of the answers from the respondents is good (0.81 – 0.90) and the quality of the items in the instrument is very good (>0.94).
- The MNSQ person infit value is 1.16, and the MNSQ person outfit is 1.06. In general, this indicates that the overall pattern of participants' answers to the instrument is good (close to 1.00).
- The value of the MNSQ infit item is 1.01, and the MNSQ outfit item is 1.06. This indicates that the overall instrument or scale is classified as good.
- The grouping of persons (participants) and items can be seen from the value of separation. The greater the value of separation, the better the quality of interaction between participants and items because it can identify differences in groups of participants and items. The person separation value of 3.23 is relatively high (> 3), while the value of item separation is equal to 4.3 (< 5). This means that the statement items can distinguish the diversity of participants, but highly variable participants can detect the items of the measuring instrument.

4.2. Item Fit Level (Item Fit)

The Rasch model analysis revealed the interaction between participants and items; participants with high ability will respond appropriately to the item's difficulty level. In this context, the parameters infit (information-weighted fit) and outfit (outliers-sensitive fit) are calculated from the mean square and standardized values, respectively. The expected value of the mean square is 1. (one). A bigger than one mean-square value for infit or outfit, say 1.3, shows that the observed data has 30% more variance than expected by Rasch. A number less than one, for example, 0.78 (1-0.22=0.78), shows that the empirical data has 22% less variance than anticipated by the Rasch model (Bond & Fox, 2015).

Some of the fit indexes available in Rasch's analysis are Person Infit ZSTD, Person Outfit ZSTD, Person Infit MNSQ, Person Outfit MNSQ, Item Infit ZSTD, Item Outfit ZSTD, Item Infit MNSQ, and Item Outfit MNSQ (Boone, Staver, & Yale, 2014).

There are several ways to find out which items do not fit::

- Comparing the MNSQ infit items with the number of Infit Mean and S.D. In this case, it is 1.01+0.48= 1.49. Based on this, it can be seen that items 11, 30, 27, and 16 have an Infit MNSQ value > 1.49 which means that these items are not appropriate or invalid.
- Viewing the MNSQ outfit value for each item. Commonly used criteria are 0.5 < MNSQ < 1.5. Referring to this criterion, it is known that items 11, 30, 27, and 16 are invalid.

4.3. Unidimensionality

The term "unidimensionality" refers to the degree to which an instrument's variety accurately measures what it is intended to measure. The table below (Table 1) summarizes the measurement results.

Table 1. Standardized residue diversity

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

	--	Empirical	Modeled
Total Raw Variance in Observations	57.6	100.0%	100.0%
Raw Variance by measures	22.6	39.2%	41.3%
Raw Variance by persons	8.7	15.1%	15.9%
Raw Variance by items	13.9	24.1%	25.4%
Raw Variance by (total)	35.0 (60.8%)	100.0%	58.7%
Unexplained variance in 1 st contrast	3.8	6.6%	10.9%
Unexplained variance in 2 nd contrast	2.9	5.0%	8.2%
Unexplained variance in 3 rd contrast	2.3	3.9%	6.4%
Unexplained variance in 4 th contrast	1.8	3.2%	5.2%
Unexplained variance in 5 th contrast	1.6	2.7%	4.5%

The study results indicate that the raw variance is 39.2 percent, which is not significantly different from the expected value of 41.3 percent. This demonstrates that the

minimum condition for unidimensionality has been met, namely raw variance greater than 20%. Additionally, the unexplained raw variance is significantly less than 15%. It

may be asserted that the instrument used to measure Indonesian Spiritual Well-being satisfies the condition for unidimensionality, which means that the statement items

4.4. Validity of the rating scale

The analysis of the validity of the rating scale is a test conducted to determine whether the choice category is confusing for participants or not. A good rating scale shows a consistent increase in each response choice in each answer category. The Indonesian Spiritual Well-being scale uses 5 alternative answers, a score of 1 (very inappropriate) to 5 (very appropriate). The results of the calculation of response ratings on the Indonesian Spiritual Well-being scale can be seen in the table 2.

A good rating scale occurs when there is consistent improvement in each response option. In the table, the average observation score starts from logit +0.16 for the

correspond to the things meant to be measured by the Indonesian Spiritual Well-being scale.

first choice, then becomes +0.14 for the second choice. This demonstrates that participants are frequently perplexed and have difficulties deciding whether to answer between responses 1 and 2. It may be shown through variable maps that participants more easily agree upon measurement items. This is demonstrated by the difference in the values of M person and M thing, where M person is more valuable. This illustration indicates that extra items are more difficult to agree on to differentiate high-ability participants. Another possibility is referring to the Anrich threshold value. The threshold value changes from zero to negative and subsequently to positive. When the Andrich threshold value is considered, it can be seen that there is a constant change of values from negative to positive, indicating that the category selection is suitable.

Table 2. Validity of the Indonesian Spiritual Well-being Scale

CATEGORY		OBSERVED		OBSVD	SAMPLE	INFIT	OUTFIT	ADRICH	CATEGORY	
LABEL	SCORE	COUNT	%	AVRGE	EXPECT	MNSQ	MNSQ	TRESHOLD	MEASURE	
1	1	849	2	.16	-.69	1.96	2.73	NONE	(-2.72)	1
2	2	2202	6	.14*	.06	1.07	1.14	-1.23	-1.29	2
3	3	9302	27	.52	.63	.85	.80	-1.09	-.09	3
4	4	11315	32	1.10	1.23	.95	.74	.73	1.26	4
5	5	11332	32	2.13	1.99	.87	.95	1.60	(2.93)	5

5. DISCUSSION

Based on the data analysis, it can be concluded that the Indonesian Spiritual Well-being scale has good psychometric qualities in general. Reliability satisfies the criteria. The advantages of Rasch's technique, as demonstrated by the dependability of person and item, were presented in this study by a high score, in the sense that participants were consistent in their responses (it was proven that there was only one outlier). Similarly, items with a high dependability score indicate that participants easily understand them.

Additionally, based on the validity review, this scale has a high unidimensionality score, indicating that all dimensions examined correspond to the Indonesian Spiritual Well-being scale. Thus, transcendence, spiritual values, and the meaning of life have been used as indicators of spiritual well being. Compared to previous instruments for measuring spiritual well being, such as those developed by Fisher or Paloudzian, this instrument incorporates local content, specifically spiritual virtues such as patience, gratitude, and sincerity. Although the initial idea of spiritual pleasure did not include any local content, the calculation of unidimensionality demonstrates that the content is an integral aspect of the scale of Indonesian Spiritual Well-being. Moreover, the item fit test revealed tat four statement items did not match the standards, necessitating their elimination. 11, 30, 27, and 16 are the four items. Participants evaluated the four things in the above statements extremely differently, resulting in a lack of

consistency in their evaluations. As a result, the number of statement items is reduced by four, leaving 31 statements.

6. CONCLUSION AND SUGGESTIONS

6.1. Conclusion

According to Rasch's modeling study, the Indonesian Spiritual Well-being scale (ISW) possesses psychometric features that match the criteria. The scale's validity standards are met by the unidimensionality score, as well as the MNSQ infit and outfit scores, indicating that the instrument is generally valid for measuring Indonesian Spiritual Well-being. According to the item appropriateness test, four statement items must be removed, leaving 31 items. Cronbach alpha scores, item, and person reliability scores are also used to meet reliability standards. However, Rasch's modeling of the ISW measuring instrument indicates that the rating scale requires additional research because the rating scale is still observed ambiguity by the partisipans. It is assumed that "social desirability bias" play an important role in supplying responses.

6.2. Suggestion

Based on the results of the Rasch modeling test, it is recommended to add statement that are considered more difficult for participants to agree with. For this reason, it is necessary to conduct a more in-depth study of the concept

of spiritual well-being in Indonesia through focus group discussions that are cross-religious, cross-cultural, and cross-educational. This is intended to obtain more diverse information so that the statement are also differentiated. In addition, more precise Rasch modeling tests can be conducted based on various demographic and geographic factors. This is meant to generate a more differentiated scale based on context, for example, different ages, given that the ISW aspect suggests a profound knowledge of life events and insights. Because diverse geographies typically correspond to distinct cultures, Rasch's concept of spiritual well being is recommended for various cultural groupings. Comparative analysis will improve the item's description or personal dependability, or it may be possible to produce a different scalogram to obtain a more representative spiritual well-being scale.

ACKNOWLEDGMENT

This research was supported by Ministry of Research, Technology, and Higher Education of the Republic of Indonesia and Research & Community Services of Universitas Tarumanagara.

REFERENCES

- [1] R. M. Ryan, V. Huta, and E. L. Deci, "Living well: A self-determination theory perspective on eudaimonia," *J. Happiness Stud.*, vol. 9, no. 1, pp. 139–170, 2008, doi: 10.1007/s10902-006-9023-4.
- [2] D. J. Disabato, F. R. Goodman, T. B. Kashdan, J. L. Short, and A. Jarden, "Different types of well-being? A cross-cultural examination of hedonic and eudaimonic well-being," *Psychol. Assess.*, vol. 28, no. 5, pp. 471–482, 2016, doi: 10.1037/pas0000209.
- [3] V. Huta and A. S. Waterman, "Eudaimonia and its distinction from hedonia: Developing a classification and terminology for understanding conceptual and operational definitions.," *J. Happiness Study*, vol. 15, pp. 1425–1456, 2014, doi: <https://doi.org/10.1007/s10902-013-9485-0>.
- [4] D. . Haybron, *Philosophy and the science of subjective wellbeing: The science of subjective wellbeing*, Eid & Lars. New York: The Guilford Press, 2008.
- [5] K. L. Pargament, *Spirituality integrated psychotherapy. Understanding and Addressing the Sacred*. New York: The Guilford Press, 2007.
- [6] R. F. Paloutzian and C. L. Park, *Handbook of psychology, religion and spirituality*. New York: The Guilford Press, 2005.
- [7] J. Fisher and N. David, "Presenting a 4 item spiritual well being index (4-ISWBI)," *Religion*, vol. 8, p. 179, 2017, doi: <https://doi.org/10.3390/rel8090179>.
- [8] Rapson, Gomez, and J. W. Fisher, "Domains of spiritual well-being and development and validation of the Spiritual Well-Being Questionnaire," *Pers. Individ. Dif.*, vol. 35, 20013, doi: 1975-91.
- [9] P. Darvyri *et al.*, "The spiritual well-being scale (swbs) in greek population of Attica," *Psychology*, vol. 5, pp. 1575–1582, 2018.
- [10] R. F. Paloutzian and C. W. Ellison, "No Title," *Loneliness, Spirit. well-being, Qual. life.*, 1982.
- [11] de J. Meezenbroek, E. B. Garssen, M. Van den Berg, A. Visser, Van Dierendonck, and W. B. Schaufeli, "Measuring spirituality as a universal human experience: A review of spirituality questionnaires," *J. Relig. Heal.*, vol. 51, pp. 336–354, 2012, doi: 10.1007/s10943-010-9376-1.
- [12] D. J. Lee, M. J. Sirgy, D. Efraty, and P. Siegel, *A study of quality of work life, spiritual well-being, and life satisfaction*, Giacalone, New York: M.E. Sharpe, 2003.
- [13] J. W. Fisher, "The four domains model: Connecting spirituality, health and wellbeing.," *J. Relig.*, vol. 2, pp. 17–22, 2011, DOI : 10.3390/rel2010017.
- [14] R. M. Furr, *Scale construction and psychometrics for social and personality psychology*. London: Sage Publication, 2011.
- [15] C. Kimberlin and A. Winterstein, "Validity and reliability of measurement instruments used in research," *Am. J. Heal. Pharm.*, vol. 65, no. 23, pp. 2276–2284, 2008, doi: DOI:10.2146/ajhp070364.
- [16] S. Suryabrata, *Metodologi penelitian*. Jakarta: Raja Grafindo Persada, 2008.
- [17] Sumintono and Widhiarso, *Aplikasi model Rasch untuk penelitian ilmu ilmu sosial*. Bandung: Tim Komunikata Publishing House, 2014.
- [18] R. K. Hambleton, F. Robin, and D. Xing, *Item response models for the analysis of educational and psychological test data*. In H. E. Tinsley, & S. D. Brown (Eds), *Handbook of applied multivariate statistics and mathematical modeling*. San Diego, CA: Academic Press, 2000.