Journal of Educational, Health and Community Psychology Vol 11, No 2, 2022 E-ISSN 2460-8467

Item Psychometric Property of Marital Quality Scale of Javanese People Using Rasch Model

Farida Agus Setiawati
Universitas Negeri Yogyakarta
farida_as@uny.ac.id

Rizki Nor Amelia
Universitas Negeri Semarang
rizkinoramelia@gmail.com

Siti Rohmah Nurhayati
Universitas Negeri Yogyakarta,
siti_rohmah@uny.ac.id

Abstract

The measurement of marital quality has been done in many countries, including Indonesia. This article described the characteristics of the item parameter endorsement index and item fit model in the data as the results of measurement of marital quality of Javanese people. The 420 couples (N = 840) from Daerah Istimewa Yogyakarta, Solo, Banyumas, and Pekalongan involved this research. The data collected were analyzed using the Rasch model with the Quest program. In general, the endorsement index moves higher as the category increases. There were four items from the dimension of relationship quality and twelve items from the dimensions of well-being quality that were misfit models. These indicated that the items did not measure the latent construct desired and need to be removed or reviewed.

Keywords: marital quality, item parameter, Rasch model, fit index, Javanese people

Received 20 February 2022/Accepted 29 May 2022

Introduction

Marital quality has been a global research focus and has become the most studied topic in marriage and family life. In many kinds of research, the measurement of marital quality uses different terms and concepts. Marital quality has many interpretations that are often equated with marital satisfaction, happiness, success, stability, attachment, adjustment, and couple relations (Arumugham, 2012; Dush et al., 2014; Fincham & Rogge, 2010; Graham et al., 2011; Knapp & Holman, 2010; Spanier, 2014). Bradbury et al., (2000) define marital quality as a global self or other reported evaluation of one's marriage or the behaviors within the union in terms of positive dimensions such as happiness, support, and satisfaction, and negative emotions such as tension and strain. Specifically, Fincham & Rogge (2010) define marital quality as reflecting each partner's subjective assessment of their relationship. Other scholars use the term marital stability, which refers to the affective and cognitive states along with the related actions that precede the termination of a relationship (Amato et al., 2007), the condition that determines if a relationship will last (Busby et al., 2009) and the desire to leave the relationship (Brown et al., 2008).

These definitions indicate that marital quality is presented holistically by assessing overall sentiments toward a couple's relationship, including commitment, satisfaction, and stability. Additionally, the many terms indicate the many different meanings (Mike & Luna, 2015; Norton, 1983; Zhang et al., 2013) concepts, and measuring tools for marital quality (Glenn et al., 2010; James, 2015).

A large body of literature has been dedicated to exploring the relationship between marital quality and various factors such as mental health, mindfulness, and forgiveness. Some findings suggest that low commitment, low satisfaction, and marital instability can be symptomatic of mental well-being (Duncan et al., 2018). Mindfulness, which includes nonjudgment, awareness, compassion, cultivating empathy, and understanding the partner's perspective and emotional regulation, may increase marital functioning (Atkinson, 2013; Kozlowski, 2013). Additionally, forgiveness was positively associated with positive marital quality and negatively related to negative marital quality (Braithwaite et al., 2011; Stafford et al., 2014). A previous study conducted by Fincham and Beach (2007) reported that forgiveness and marital quality appear to have a bidirectional link. Karremans et al. (2011) explored couple relationships and demonstrated that the closer an individual felt to their partner, the more likely they were to forgive them. Related to the relationship between husbands and wives, Proulx et al. (Proulx et al., 2007) and Robles et al. (Robles et al., 2014) confirmed that high-quality marriages might boost the health and well-being of couples, while low-quality marriages may harm the relationship.

Several examples of instruments have been successfully developed and used for measuring marital quality through self-report. They are the Locke–Wallace Marital Adjustment Test (LWMAT), the Kansas Marital Satisfaction Scale (KMS), the Quality of Marriage Index (QMI), the Relationship Assessment Scale (RAS), the Marital Opinion Questionnaire (MOQ), Karney and Bradbury's (1997) semantic differential scale (SMD), and the Couples Satisfaction Index (CSI) (Graham et al., 2011), as well as Relationship Quality (RQ) (Chonody et al., 2018). Meanwhile, in the Indonesian context, there are several instruments for measuring the marital quality of Indonesian people who have been successfully developed and have psychometric properties in the form of a good reliability coefficient.

Sunarti, Tati, Atat, Noorhaisma, and Lembayung (2005) used the theory of Conger et al. (1990) to

develop an instrument that measures marital quality, which consists of two dimensions, namely marital satisfaction and happiness (Sunarti et al., 2005). The two dimensions have the following indicators: commitment, trust, marital values, communication, family togetherness, equality, relationships with extended family, expressions of affection, love and sex, common interests, and family economy and income. Overall, the developed instrument has a Cronbach's alpha coefficient of 0.774.

Rumondor (2013) combines three measuring tools, namely the Dyadic Adjustment Scale (Spanier, 1976), Enrich Marital Satisfaction (Fowers & Olson, 1993), and the Marriage Satisfaction Questionnaire (Sadarjoen, 2004), to develop an instrument that functions to measure marital satisfaction in urban couples (Rumondor, 2013). The instrument developed has four alternative answers, using the Likert scale, with 58 statement items distributed into nine measuring dimensions. Finally, the overall instrument has a good alpha Cronbach (α) coefficient of 0.967, where the Cronbach alpha coefficients for each dimension are as follows. Communication (α = 0.724), role balance (α = 0.818), agreement (α = 0.794), openness (α = 0.830), intimacy (α = 0.817), social intimacy in relationships (α = 0.773), sexuality (α = 0.734), financial (α = 0.822), and spirituality (α = 0.924).

Istiqomah & Mukhlis (2015) developed an instrument to measure marital satisfaction by modifying the marital satisfaction scale of *Enrich Marital Satisfaction* (Fowers & Olson, 1993) (Istiqomah & Mukhlis, 2015). The scale has five alternative answers and initially contains 40 items, but in the end, only 24 items are valid with a Cronbach alpha coefficient of 0.845. Soraiya, Khairani, Rachmatan, Sari, & Sulistyani (2016) used Hazan & Shaver's theory (1987) to develop an instrument that measures marital attachment. The result is that the secure attachment has α = 0.864; avoidance attachment has α = 0.877; and anxiety attachment has α = 0.691 (Soraiya et al., 2016).

The various examples of the marital quality instruments show that reliability is the focus of psychometric property analysis. Whereas reliability estimated using the Classical Test Theory approach (e.g., Cronbach Alpha) found weaknesses that depend on the particular samples of individuals who responded to the instrument (Auné et al., 2020; Rusch et al., 2017) and affect the accuracy and objectivity of the resulting measurements (AL-khadher & Albursan, 2017). The reliability

coefficient of a measuring instrument only applies to the research sample because it depends on the data sample. The magnitude of the coefficient will be different in other research samples, so the reliability coefficient will always be estimated for each research data, even though the instrument is the same. Besides, applying a conventional Cronbach Alpha is more appropriate to estimate the reliability of a unidimensional construct instrument. If multidimensional, it will underestimate the true reliability (Widhiarso & Ravand, 2014). Although, the Alpha coefficient is often used to estimate the reliability of the non-cognitive measurement. In the actual, a psychological instrument is more multidimensional than a unidimensional construct

The modern test theory is based on invariance scores, that is not dependent on the sample research. The Rasch model is part of the Modern Test Theory (Andrich, 2011). It is like one logistic parameter in items theory response (IRT). The models of IRT have been classified based on the categorization of the response answer. If the response is dichotomous, the models are one logistic parameter (1-PLM), two logistic parameters (2-PLM), or three logistic parameters (3-PLM). If the response categories are more than two, the polytomous IRT is the most suitable model (Auné et al., 2020).

Much of the literature asserts that the 1-PL IRT model is the same as the Rasch model. This opinion is misleading. The mathematical equations can look similar, but their motivations are entirely different (Linacre, 2012), which means that the Rasch model is not part of IRT (Hayat et al., 2020). The Rasch Rating Model (Andrich, 1978; Masters, 1982) is the most widely used for ordinal variables that are assumed to collectively reflect common latent variables (Adams et al., 2012).

Andrich Rating Scale Model (RSM) is a member of the Rasch model when the response is ordinal, like the Likert scale (Andrich, 2016). The model fasilitated item with a stem (statement of attitude) and the respondent is required to mark a response on the disagree to agree, indicating the extent to which the statement in the stem is endorsed (Bond & Fox, 2015). The response strongly disagree to strongly agree will be scored continuum 0,1,2,3. It the easier if the endorsement index is defined as the transition point of a response from one category to an adjacent variety on a Likert scale (Putra & Retnawati, 2020). The endorsement index is a boundary location parameter on a continuum between the k and k-1 categories of a scale (Gómez et al., 2012). For example, response strongly disagree to disagree

(0 to 1), disagree to agree (1to 2) and agree to strongly agree (2 to 3). With the estimation of the endorsement index, it becomes clear that the items do not have the same relative value in the construct being examined (Teman, 2013). In this study, psychometric properties are revealed by a modern test theory approach using the Rasch model. The polytomous data such as the Likert-type scale, the analysis method specifically takes the form of the Andrich Rating Scale Model (Andrich RSM) that a rating response mechanism for ordered categories.

Research studies to detect psychometric properties of measuring instruments using RSM have been studied in many countries. In the USA, RSM was applied to 25 items in the Evaluate Patient-Reported Outcome measurement analyzed using RSM, and got 25 fit items and several responses that have the potential to be problematic (Petrillo et al., 2015). Analysis of RSM was also carried out to measure undergraduate student anxiety with several items that must be re-evaluated (Teman, 2013). Forty-three items Iranian Voice Quality of Life Profile (IVQLP) are fit by the RSM model (Dehqan et al., 2017). Rasch rating scale model has been examined and reassessed the psychometric properties of the Persian version of the PedsQLTM 4.0 Generic Core Scales in children (Jafari et al., 2012). Six items of The Proactive Personality Scale were tested on students in Ghana, and got the results of a fit model (Teye-kwadjo & Bruin, 2021). Application of the rating scale model also used in the assessment of the quality of life of Spain persons with intellectual disability (Gómez et al., 2012).

These studies show that RSM has been widely studied in various countries. The research of psychometric properties by the Andrich RSM on the marital quality scale has not been done in the original Indonesian version. Therefore, this study aims to describe the psychometric properties of item parameters using the Rasch model on the Javanese marital quality measurement scale.

Method

Participant

The research sample was 420 married couples (N = 840) who identified themselves as Javanese people. All subjects live in Java Island, which is represented by the Daerah Istimewa Yogyakarta, Solo, Banyumas, and Pekalongan. Data collection technique using multistage random sampling. the first random is based on the district in the four regions. The next random one is based on

the village in the district.

Measurement

The research instrument used was a marital quality scale consisting of 57 statement items in a Likert-type with five alternative response choices; absolutely inappropriate, not appropriate, somewhat appropriate, appropriate, and absolutely appropriate. The construct of the Javanese marital quality scale consists of two dimensions, namely the dimension of Relation Quality represented by items 1 to 39, and the dimension of Well-being Quality represented by items 40 to 57. The measurement of the response model is a manifestation of a person's latent dimension or factor. It reflected a single continuum to represent the latent variable. So the requirements for analysis with the Rasch model are unidimensional. Because this instrument has two dimensions, the instrument analysis is carried out by separating the dimensions. It, is relevant to the assumption of the Rasch model is unidimensionality.. The evidence of unidimensionality from factor analysis of the data using the principal component of Confirmatory Factor Analysis (CFA). It was done to test the construct dimensionality of data from the marital quality instrument. The result of CFA has a good fit model with two dimensions (RMSEA = .075; GFI = .97; AGFI = .93; NFI = .99; CFI = .99; IFI = .99),

Data collected and analyzed.

The data were collected in July 2019. Before completing the questionnaire, respondents filled out an informed consent form to participate in this study. Respondents also received a token of gratitude from the researcher. The respondent got the questionnaire from assistant research and finished off all items. After all of the data were collected, they were analyzed using the Rasch Model (Andrich RSM) with the help of the QUEST program. This program was chosen because it can display the endorsement index value in the form of categories for each item analyzed. The Andrich RSM is an appropriate analysis method for rating data like the Likert type (Ambiel et al., 2015; Linacre, 2000). Andrich's formulation for the Rating Scale Model is presented in equation (1). In that function, n and i are the locations of person n on item i, respectively, j is the location of the jth step in each item, and k is the category (Andrich, 1978; De Ayala, 2009).

$$P_{xni} = \frac{exp\left[\sum_{j=0}^{x} \tau_j + x(\theta_n - \delta_i)\right]}{\sum_{k=0}^{m} exp\left[\sum_{j=0}^{x} \tau_j + x(\theta_n - \delta_i)\right]} \ldots\ldots (1)$$

Exp = exponential (e)= 1,7[a]
n = person n
i = item i
k = lowest category=0
m = highest catogory
τ_j = category coefficient
θ_n = person location (marital quality person n)
δ_i =item location (indorsing the item)

Result

Psychometric analysis of the Javanese marital quality scale in this study used the Modern Test Theory approach, namely the Rasch model, which only contains one item parameter, namely the item parameter location. This parameter in non-cognitive instruments is called the probability of endorsement or endorsement index, while in cognitive test instruments, it is called the difficulty index or threshold. Table 1 is the index endorsement of marital quality. There are four endorsement indexes in each item, which come from five categories on the Likert scale. P-1 is a characteristic in categories 1 and 2, P-2 is a characteristic in categories 2 and 3, P-3 is in categories 3 and 4, and P-4 is a characteristic in categories 4 and 5. The all-endorsement index increases as the category increases, both for the items on the dimension of relationship quality and well-being quality.

Table 1
Endorsement indexes the items of marital quality

Dimension	Item	Endorsement Index Category (P-i)			
		1	2	3	4
Relation Quality	1	-1.59	-.75	-.10	3.13
	2	-6.50	-.55	.23	3.51
	3	-2.38	-.79	.29	3.06
	4	-1.88	-1.24	-.34	2.89
	5	-2.75	-1.22	.15	3.37
	6	-6.44	-.80	.40	3.67
	7	-1.81	-1.22	-.62	2.27
	8	-6.00	-1.52	-.70	2.50
	9	-1.88	-1.22	-.50	2.95
	10	-2.31	-1.24	-.94	1.44
	11	-2.42	-1.49	-.26	2.42
	12	-1.44	-1.25	-.64	2.36
	13	-1.73	-1.22	-.75	1.91
	14	-1.38	-.84	.05	2.94
	15	-1.63	-.91	.27	3.08
	29	-1.06	-.64	.12	2.98
	30	-1.34	-.58	.53	3.65
	18	-1.09	.47	1.53	3.50
	19	-.72	-.31	.31	2.53
	20	-1.55	-.54	.06	2.60
	21	-1.19	.23	1.11	3.32
	22	-1.98	-1.75	-1.16	1.92
	23	-1.50	-.13	1.05	3.87
	24	-1.31	-.86	-.11	2.55
	25	-1.66	-.90	-.45	2.62
	26	-1.59	-1.09	.41	3.00
	27	-1.28	-0.83	-.24	2.62
	28	-1.30	-1.08	-.53	2.40
	31	-1.91	-.26	.64	2.77
	32	-2.28	-.80	-.06	2.49
	33	-1.03	-.22	.63	2.88
	34	-1.13	-.73	.00	3.47
	35	-1.05	0.69	-.18	3.01
	36	-2.05	-1.15	-.75	2.94
	37	-2.13	-.90	-.31	2.39
	38	-1.94	-1.08	-.46	2.48
	39	-1.95	-1.08	-.44	2.20
Well-being Quality	40	-2.00	-.07	1.41	5.02
	41	-2.22	-.18	1.51	5.57
	42	-1.84	-.63	.92	4.46
	43	-2.25	-.65	1.18	5.07
	44	-1.92	-1.50	-.20	4.13
	45	-2.13	-.94	.28	4.23
	46	-2.25	-1.55	-.65	2.35

Table 1
Endorsement index the items of marital quality (continue)

Dimensian	Items	Endorsement Index Category (P-i)			
		1	2	3	4
Well-being	47	-1.03	.03	.60	3.43
Quality	48	-2.25	-1.55	-.65	2.52
	49	-2.09	-1.64	-.81	2.81
	50	-1.45	-1.29	-.89	2.34
	51	-2.31	-1.47	-.65	2.59
	52	-1.69	-1.69	-1.51	1.54
	53	-1.73	-1.73	-1.43	1,70
	54	-2.00	-1.49	-1.12	1.59
	55	-1.83	-1.83	-1.21	2.03
	56	-1.75	-.87	.41	3.61
	57	-2.00	-1.69	-.93	2.24

Other information that can be extracted from the psychometric property analysis using the QUEST program is the item fit model. The items fit against the Rasch Model is identified by whether the items cross the boundary line of the fit index set by the program or not. If the item crosses the right boundary line, then the item is called underfit, whereas if the item crosses the left boundary line, then the item is called overfitting. Distributing the fit items presented in Figures 1 and 2. Based on the range score ranging from 0.7 to 1.2, it can be seen that there were four items (number 10, 12, 16, and 18) from the dimension of Relation Quality, and from figure two, twelve items (number 40, 41, 46, 47, 48, 49, 50, 51, 52, 53, 55, and 57) from the dimensions of Well-being Quality that were misfit model. There are 10.26% of items that are misfit (5.13% items are overfit and 5.13% items are underfit) in the dimension of Relation Quality, and 66.67% items are misfit (50% items are overfit, and 16.67% items are underfit) in the dimension of Well-being Quality.

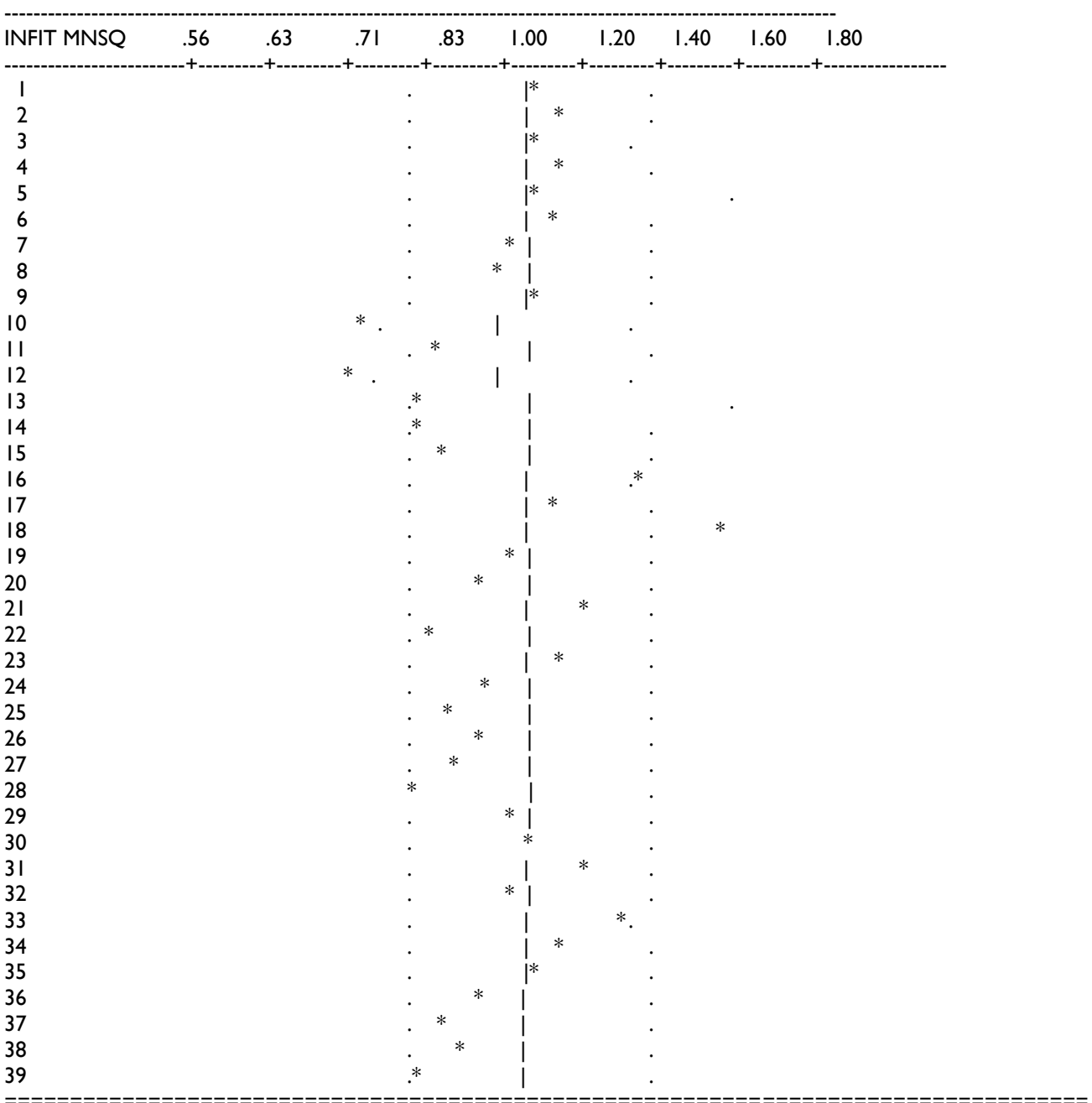

Figure 1. INFIT MNSQ Map of Relation Quality Dimention

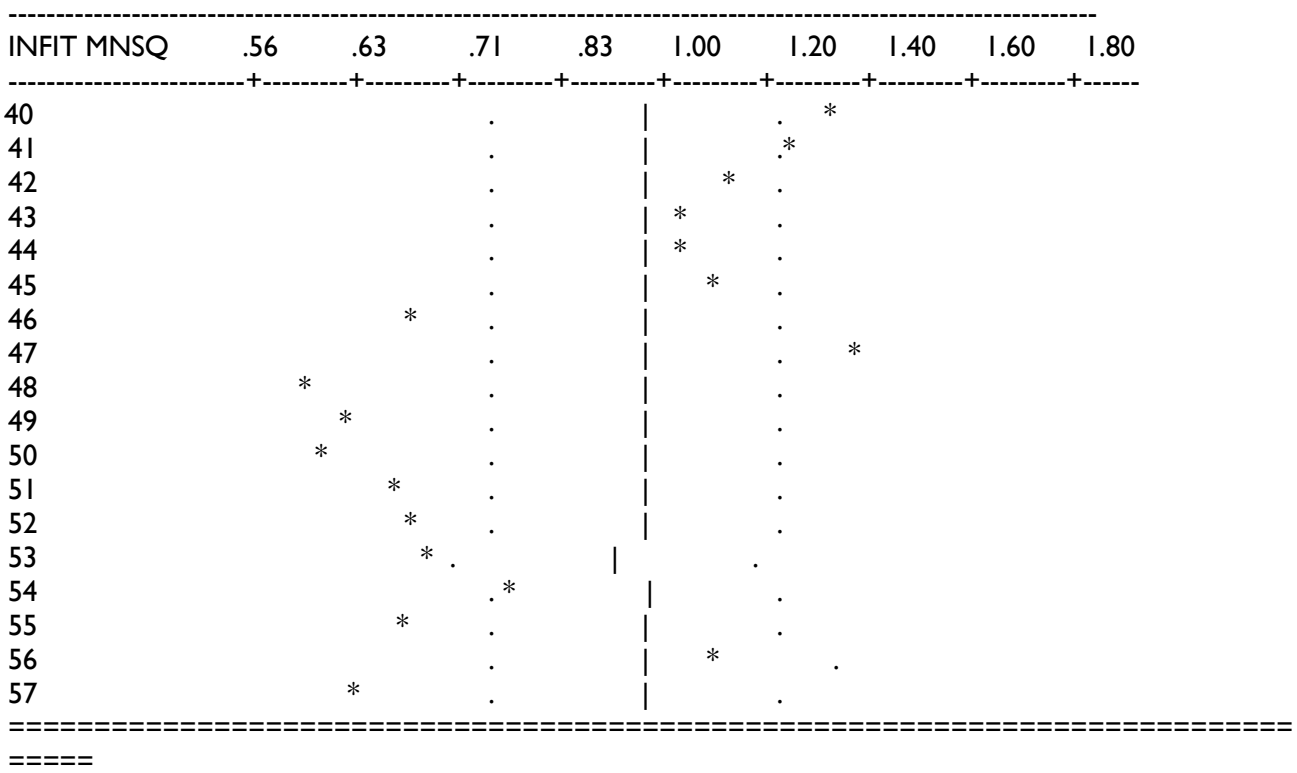

Figure 2. INFIT MNSQ Map of Well-being Quality Dimension

Discussion

Psychometric property analysis is the fundamental for test developer to be carried out on a scale of measurement developed, mainly using the Modern Test Theory approach, sometimes called Item Response Theory or IRT (Ashraf & Jaseem, 2020; Magno, 2009). It is because the psychometric properties in this modern analysis are assumed to be independent samples or parameter invariance (Gomez & Fisher, 2005). In addition, parameters are also provided at the item level, this approach would allow for the identification of items that are functioning differently in terms of their ability to discriminate, represent and reliably measure the traits at different levels of the underlying trait; this, in turn, can facilitate the development and revision of measures (Gomez & Fisher, 2005). Thus, the use of Modern Test Theory will provide more valuable data about the psychometric properties of the scale and its items and provide useful directions for their improvement.

The study found the endorsement index moves higher as the category increases, both for Relation Quality and the dimensions of Well-being Quality. Further, item 35 (even though it is difficult, my partner and I are trying to carry out our respective obligations) is one of the items measuring the Cooperation aspect of the dimension of the Relation Quality, indicating that this item has a reversed endorsement index value for category two and three ($\tau_2 = .69$ logit and $\tau_3 = -.18$ logit), whereas the endorsement index value should increase monotonically from the lowest rating point to the highest rating point (Bond & Fox, 2015). If the category functions properly, then the endorsement index value will be in the same order as the category (Houghton et al., 2017). Due to having a reversed endorsement index value, it can be concluded that category two and three in item 35 do not function properly. The category that has a reversed endorsement index value gives an indication of the low frequency of responses to that category by respondents (Adams et al., 2012).

There are a several of reasons why this can be a problem, and whenever it does, it should be immediately reviewed carefully by the scale constructor (Adams et al., 2012). Therefore, if a disordered threshold occurs, modification is needed, for example by combining the reversed categories with the previous category (collapsing categories) (Adams et al., 2012; Houghton et al., 2017). Other findings are item 52 (I thank God for my marriage) and item 53 (I am grateful for my marriage) which are items measuring the aspect of Happiness on the dimension of Well-being Quality which has the same endorsement index value in the category one and two (Item 52: $\tau_1 = -1.69$ logit; and $\tau_2 = -1.69$ logit; Item 53: $\tau_1 = -1.73$ logit; and $\tau_2 = -1.73$ logit). The meaning is the same as item 35 that the category does not function properly, so the solution also needs to be modified.

In addition to these three items, all the marital quality Javanese people scale items already have an endorsement index value by the rules. Good items will have an endorsement index value that increases as the category increases, which means that respondents who have higher abilities will support progressively higher categories, and vice versa (Bond & Fox, 2015). For example, on item one (my partner and I exchange ideas in many ways) which is one of the items measuring the Communication aspect of the Relation Quality dimension, the endorsement index value increases as

the category increases ($\tau_1 = -1.59;\ \tau_2 = -.75;\ \tau_3 = -.10;\ \tau_4 = 3.13$). In category 1, the endorsement index value is -1.59 logit, this means that the average estimated ability for all respondents who choose category one on item one is -1.59 logit. In category two, the endorsement index value is -.75 logit, this means that the average Other information that can be extracted from the psychometric property analysis using the QUEST program is the item fit model. The items fit against the Rasch Model is identified by whether the items cross the boundary line of the fit index set by the program or not. If the item crosses the right boundary line, then the item is called underfit, whereas if the item crosses the left boundary line, then the item is called overfitting. Distributing the fit items presented in Figures 1 and 2. Based on the range score ranging from 0.7 to 1.2, it can be seen that there were four items (number 10, 12, 16, and 18) from the dimension of Relation Quality, and from figure two, twelve items (number 40, 41, 46, 47, 48, 49, 50, 51, 52, 53, 55, and 57) from the dimensions of Well-being Quality that were misfit model. There are 10.26% of items that are misfit (5.13% items are overfit and 5.13% items are underfit) in the dimension of Relation Quality, and 66.67% items are misfit (50% items are overfit, and 16.67% items are underfit) in the dimension of Well-being Quality estimated ability for all respondents who choose category two on item one is -.75 logit. In category three, the endorsement index value is -.10 logit, this means that the average estimated ability for all respondents who choose category three on item one is -.10 logit. In category four, the threshold value is 3.13 logit, this means that the average estimated ability for all respondents who choose category four on item one is 3.13 logit.

Items fit related and represent the degree of difference between the observed responses to the article and the model expected (Kean et al., 2018). The misfitting items indicated that a person's ability or item difficulty does not contribute to the underlying test construct (Bayne & Hankey, 2020; Tavakol & Dennick, 2013). In Figure 1 and Figure 2, the data presented the item fit model in each dimension. In the QUEST program, the item fit criteria used are INFIT MNSQ (Inlier-sensitive or Information-weighted Fit Mean-Square), The vertical dotted line indicates the range of INFIT MNSQ values received, between 0.77 – 1.30 (Adams & Khoo, 1998). INFIT are statistics that describe the sensitivity patterns of response to items targeted on the person, and vice versa (Linacre, 2002). The condition is called overfit when the INFIT value is lower than the referenced bound (the data condition is too ideal and predictable). When the INFIT value is greater than the referenced bound,

the state is called underfit. The data condition is inconsistent and difficult to predict (Linacre, 2012). Based on Figure 1 and Figure 2, it can be seen that there are four misfit items in the dimensions of Relation Quality and 12 misfit items in the dimension of Well-being Quality. The details of the misfit items are as follows. The Items in the dimension of Relation Quality are: the item 10 (my partner and I love each other), item 12 (my partner and I try to make each other happy), and two overfit items, namely item 16 (my partner and I usually eat together) and item 18 (my partner and I still take the time to go together). In the dimension of Well-being, nine underfit items were detected: item 46 (I feel comfortable living with my partner), item 48 (I feel peaceful in my marriage), item 49 (I feel comfortable living with my partner), item 50 (I enjoy my married life), item 51 (I feel comfortable in my household), item 52 (I thank God for my marriage), item 53 (I am grateful for my marriage), item 55 (I am happy with my marriage), and item 57 (I feel happy living with my partner). Three overfit items, namely item 40 (my partner and I rarely argue), item 41 (there are rarely disputes between my partner and me), and item 47 (I never worry about my household). The misfit items indicate that they did not really measure the latent construct desired (Ishar & Roslin, 2016), so they can be removed (Reid et al., 2007) or reviewed (Mutalib et al., 2015).

This study successfully described the characteristics of the item parameter endorsement index and item fit model in the data as the results of the measurement of marital quality of Javanese people using the Modern Test Theory. Still, this study has some limitations, such as the analysis results found several items that did not infit, especially the well-being items. These results are obtained from the Rasch model analysis, which only considers one parameter, namely item location. The use of one parameter makes other parameters, such as item discrimination, not considered in the estimation of this model. Using different methods in analyzing items with more than one logistic parameter can be done to see the consistency of the results of the fit items. The results of the item analysis of this instrument are limited to the marital quality instrument, which the researcher developed with the construct of relation and well-being dimensions. Efforts to look at different dimensions such as unidimensionality or more than two dimensions allow for better fit model results.

Conclusion

Psychometric analysis of the Javanese marital quality scale using the Rasch model contains one item

parameter called the probability of endorsement or endorsement index. These index values move higher as the categories. Forty-one items fit against the Rasch Model. Four items (number 10, 12, 16, and 18) from the dimension of Relation Quality and 12 items (number 40, 41, 46, 47, 48, 49, 50, 51, 52, 53, 55, and 57) from the dimensions of Well-being Quality that were misfit model. These indicated that several items did not measure the latent construct. Thus, further development of measuring instruments is needed to obtain a more accurate measurement of the marital quality of Javanese people. The fit items in this study can be a recommendation to measure marital quality. Computer-based testing for the administration process is easier and more precise to use. Furthermore, the benefit of marital quality measurements can be used personally for self-evaluation in improving the quality of marriage and institutions that deal with marital and family problems.

Acknowledgments

The article from multy years research was funded by Direktorat Jenderal Pendidikan Tinggi, Kementerian Pendidikan dan Kebudayaan, Riset, dan Teknologi, Indonesia (grand number 71/Penelitian/PD/UN34.21/2019 and 046/F4.1/AK.04.PT/2021)

References

Adams, & Khoo, S.-T. (1998). ACER Quest: The interactive test analysis system. In *Australian Council for Educational Research*. The Australian Council for Educational Research.

Adams, R. J., Wu, M. L., & Wilson, M. (2012). The rasch rating model and the disordered threshold controversy. *Educational and Psychological Measurement*, *72*(4), 547–573. https://doi.org/10.1177/0013164411432166

AL-khadher, M. M. A., & Albursan, I. S. (2017). Accuracy of measurement in the classical and the modern test theory: An empirical study on a children intelligence test. *International Journal of Psychological Studies*, *9*(1), 71–80. https://doi.org/10.5539/ijps.v9n1p71

Amato, P. R., Booth, A., Johnson, D. R., & Rogers, S. J. (2007). *Alone together: How marriage in America is changing*. Harvard University Press.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*(4), 561–573. https://doi.org/10.1007/BF02293814

Andrich, D. (2011). Rating scales and Rasch measurement. *Expert Review of Pharmacoeconomics & Outcomes Research*, *11*(5), 571–585. https://doi.org/10.1586/erp.11.59

Andrich, D. (2016). *Rasch Rating-Scale Model* (1st editio, p. 20). Chapman and Hall/CRC.

Arumugham, P. (2012). Marital quality in inter-caste and intra-caste marriages. *International Journal of Interdisciplinary Social Sciences*, *6*(9), 17–26. https://doi.org/10.18848/1833-

1882/CGP/v06i09/52155

Ashraf, Z. A., & Jaseem, K. (2020). Classical and modern methods in item analysis of test tools. *International Journal of Research and Review*, *7*(5), 397–403.

Atkinson, B. J. (2013). *Mindfulness training and the cultivation of secure, satisfying couple relationships*. *2*(2), 73–94. https://doi.org/10.1037/cfp0000002

Auné, S. E., Abal, F. J. P., & Attorresi, H. F. (2020). Análisis psicométrico mediante la Teoría de la Respuesta al Ítem: modelización paso a paso de una Escala de Soledad. *Ciencias Psicológicas*, *14*(1), 1–15. https://doi.org/10.22235/cp.v14i1.2179

Bayne, H. B., & Hankey, M. S. (2020). Development and Rasch analysis of the empathic counselor response scale. *Measurement and Evaluation in Counseling and Development*, *53*(3), 182–194. https://doi.org/10.1080/07481756.2019.1691462

Bond, T. G., & Fox, C. M. (2015). Applying the Rasch model: Fundamental mesurement in the human sciences. In *Applying the Rasch Model* (Third ed). Routledge.

Bradbury, T. N., Fincham, F. D., & Beach, S. R. H. (2000). *Research on the nature and determinants of marital satisfaction: A decade in review*. *62*(November), 964–980.

Braithwaite, S. R., Selby, E. A., & Fincham, F. D. (2011). *Forgiveness and relationship satisfaction: Mediating mechanisms*. *25*(4), 551–559. https://doi.org/10.1037/a0024526

Brown, E., Orbuch, T. L., & Bauermeister, J. A. (2008). Religiosity and marital stability among black american and white american couples. *Family Relations*, *57*(2), 186–197. https://doi.org/https://doi.org/10.1111/j.1741-3729.2008.00493.x

Busby, D. M., Holman, T. B., & Niehuis, S. (2009). *The association between partner enhancement and self-enhancement and relationship quality outcomes*. *71*(August), 449–464.

Chonody, J. M., Gabb, J., Killian, M., & Dunk-West, P. (2018). Measuring relationship quality in an international study: Exploratory and confirmatory factor validity. *Research on Social Work Practice*, *28*(8), 920–930. https://doi.org/10.1177/1049731516631120

De Ayala, R. J. (2009). The theory and practice of Item Response Theory. In *The Guildford Press*. The Guildford Press.

Dehqan, A., Fariba Yadegari, Asgari, A., C.Scherer, R., & Dabirmoghadam, P. (2017). Development and validation of an Iranian Voice Quality of Life Profile (IVQLP) based on a Classic and Rasch Rating Scale Model (RSM). *Journal of Voice*, *31*(1), 113.e19-113.e29.

Dush, C. M. K., Taylor, M. G., & Kroeger, R. a. (2014). Marital happiness and psychological well-being across the life course *. *Family Relations*, *57*(2), 211–226.

Fincham, F. D., & Rogge, R. (2010). Understanding relationship quality: Theoretical challenges and new tools for assessment. *Journal of Family Theory & Review*, *2*(4), 227–242. https://doi.org/10.1111/j.1756-2589.2010.00059.x

Glenn, N. D., Uecker, J. E., & Love, R. W. B. (2010). Later first marriage and marital success. *Social Science Research*, *39*(5), 787–800. https://doi.org/10.1016/j.ssresearch.2010.06.002

Gómez, L. E., Arias, B., Verdugo, M. Á., & Navas, P. (2012). Application of the Rasch rating scale model to the assessment of quality of life of persons with intellectual disability. *Journal of*

Intellectual & Developmental Disability, *37*(2), 141–150. https://doi.org/10.3109/13668250.2012.682647

Gomez, R., & Fisher, J. W. (2005). Item response theory analysis of the spiritual well-being questionnaire. *Personality and Individual Differences*, *38*(5), 1107–1121. https://doi.org/10.1016/j.paid.2004.07.009

Graham, J. M., Diebels, K. J., & Barnow, Z. B. (2011). The reliability of relationship satisfaction: A reliability generalization meta-analysis. *Journal of Family Psychology*, *25*(1), 39–48. https://doi.org/10.1037/a0022441

Hayat, B., Putra, M. D. K., & Suryadi, B. (2020). Comparing item parameter estimates and fit statistics of the Rasch model from three different traditions. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *24*(1), 39–50. https://doi.org/10.21831/pep.v24i1.29871

Houghton, S., Wood, L., Marais, I., Rosenberg, M., Ferguson, R., & Pettigrew, S. (2017). Positive mental well-being: A validation of a rasch-derived version of the warwick-edinburg mental well-being scale. *Assessment*, *24*(3), 371–386. https://doi.org/10.1177/1073191115609995

Ishar, N. I. M., & Roslin, R. M. (2016). Customer voice retaliation (CVR) construct verification: A rasch analysis approach. *Procedia Economics and Finance*, *37*(16), 214–220. https://doi.org/10.1016/S2212-5671(16)30116-2

Istiqomah, I., & Mukhlis. (2015). Hubungan antara religiusitas dengan kepuasan kerkawinan. *Jurnal Psikologi*, *11*(2), 71–78.

Jafari, P., Bagheri, Z., Ayatollahi, S. M. T., & Soltani, Z. (2012). Using Rasch rating scale model to reassess the psychometric properties of the Persian version of the PedsQL TM4.0 Generic Core Scales in school children. *Health and Quality of Life Outcomes*, *10*(1), 27. https://doi.org/10.1186/1477-7525-10-27

James, S. L. (2015). Variation in trajectories of women's marital quality. *Social Science Research*, *49*(Januari), 16–30. https://doi.org/10.1016/j.ssresearch.2014.07.010

Karremans, J. C., Regalia, C., Paleari, F. G., Fincham, F. D., Cui, M., Takada, N., Ohbuchi, K., Terzino, K., Cross, S. E., & Uskul, A. K. (2011). *Maintaining harmony across the globe : The cross-cultural association between closeness and interpersonal forgiveness*. https://doi.org/10.1177/1948550610396957

Kean, J., Bisson, E. F., Brodke, D. S., Biber, J., & Gross, P. H. (2018). An introduction to item response theory and rasch analysis: Application using the eating assessment tool (EAT-10). *Brain Impairment*, *19*(1), 91–102. https://doi.org/10.1017/BrImp.2017.31

Knapp, S. J., & Holman, T. B. (2010). Introducing a special issue: on the need to theorize marital quality. *Journal of Family Theory & Review*, *2*(4), 221–226. https://doi.org/10.1111/j.1756-2589.2010.00058.x

Kozlowski, A. (2013). *Sexual and relationship therapy mindful mating : exploring the connection between mindfulness and relationship satisfaction*. *October 2014*, 37–41. https://doi.org/10.1080/14681994.2012.748889

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch*

Measurement Transactions, *16*(2), 878.

Linacre, J. M. (2012). *Winsteps help for rasch analysis*.

Magno, C. (2009). Demonstrating the difference between classical test theory and item response theory using derived test data. *The International Journal of Educational and Psychological Assessment*, *1*(1), 1–11.

Mike, P. G., & Luna, J. (2015). *An approach to the marital quality and the factors*. *2*(1), 23–32.

Mutalib, A. A., Ghafar, H. A., Baharom, S., & Hamzah, N. (2015). Rasch model analysis of implementation effectiveness for final year research project course in civil & structural engineering department, UKM. *Journal of Engineering Science and Technology*, *10*(June), 11–22.

Norton, R. (1983). Measuring marital quality: a critical Look at the dependent variable. *Journal of Marriage and the Family*, *45*(3), 464. https://doi.org/10.2307/351652

Petrillo, J., Cano, S. J., McLeod, L. D., & Coon, C. D. (2015). Using classical test theory, item response theory, and rasch measurement theory to evaluate patient-reported outcome measures: A comparison of worked examples. *Value in Health*, *18*(1), 25–34. https://doi.org/10.1016/j.jval.2014.10.005

Proulx, C. M., Helms, H. M., & Buehler, C. (2007). Marital quality and personal well-being: A meta-analysis. *Journal of Marriage and Family*, *69*(3), 576–593. https://doi.org/https://doi.org/10.1111/j.1741-3737.2007.00393.x

Putra, M. D. K., & Retnawati, H. (2020). Rasch analysis of an indonesian version of the general self-efficacy scale-12: A comparison of rating scale model (RSM) and partial credit model (PCM). *Indonesian Journal of Educational Assesment*, *3*(1), 52. https://doi.org/10.26499/ijea.v3i1.55

Reid, C. A., Kolakowsky-Hayner, S. A., Lewis, A. N., & Armstrong, A. J. (2007). Modern psychometric methodology: Applications of item response theory. *Rehabilitation Counseling Bulletin*, *50*(3), 177–188. https://doi.org/10.1177/00343552070500030501

Robles, T. F., Slatcher, R. B., Trombello, J. M., & Mcginn, M. M. (2014). *Marital quality and health : A meta-analytic review*. *140*(I), 140–187.

Rumondor, P. C. . (2013). Pengembangan alat ukut kepuasan pernikahan pasangan urban. *Humaniora*, *4*(2), 1134–1140.

Rusch, T., Lowry, P. B., Mair, P., & Treiblmaier, H. (2017). Breaking free from the limitations of classical test theory: Developing and measuring information systems scales using item response theory. *Information & Management*, *54*(2), 189–203. https://doi.org/10.1016/j.im.2016.06.005

Soraiya, P., Khairani, M., Rachmatan, R., Sari, K., & Sulistyani, A. (2016). Kelekatan dan kepuasan pernikahan pada dewasa awal di kota Banda Aceh. *Jurnal Psikologi Undip*, *15*(1), 36. https://doi.org/10.14710/jpu.15.1.36-42

Spanier, G. B. (2014). *Marital quality : A review of the seventies *. *42*(4), 825–839.

Stafford, L., David, P., & Mcpherson, S. (2014). *Sanctity of marriage and marital quality*. *31*(1), 54–70. https://doi.org/10.1177/0265407513486975

Sunarti, E., Tati, A. S. N., Noorhaisma, R., & Lembayung, D. P. (2005). Pengaruh tekanan ekonomi keluarga, dukungan sosial, kualitas perkawinan, pengasuhan, dan kecerdasan emosi anak

terhadap prestasi belajar anak. *Media Gizi & Keluarga*, *1*(29), 34–40.

Tavakol, M., & Dennick, R. (2013). Psychometric evaluation of a knowledge based examination using Rasch analysis: An illustrative guide: AMEE Guide No. 72. *Medical Teacher*, *35*(1), e838–e848. https://doi.org/10.3109/0142159X.2012.737488

Teman, E. D. (2013). A rasch analysis of the statistical anxiety rating scale. *Journal of Applied Measurement*, *14*(4), 414–434.

Teye-kwadjo, E., & Bruin, G. P. De. (2021). *Rasch Analysis of the Proactive Personality Scale*. 1–19. https://doi.org/10.1177/00332941211028110

Widhiarso, W., & Ravand, H. (2014). Estimating reliability coefficient for multidimensional measures: A pedagogical illustration. *Review of Psychology*, *21*(2), 111–121.

Zhang, H., Xu, X., & Tsang, S. K. M. (2013). Conceptualizing and validating marital quality in Beijing: A pilot study. *Social Indicators Research*, *113*(1), 197–212. https://doi.org/10.1007/s11205-012-0089-6