Psychometric Properties of Character Strength Instruments Using the Rasch Model

Yulinda Erma Suryani[1,2], Farida Agus Setyowati[3], Tukiyo[1], and Adila Eka Saputri[1]
[1]*Psychology Department Universitas Widya Dharma Klaten, Jawa Tengah, Indonesia*
[2]*Educational Research and Evaluation, Doctoral Program, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia*
[3]*Psychology Department Universitas Negeri Yogyakarta, Yogyakarta, Indonesia*
yulindaerma.2021@student.uny.ac.id

Keywords: Character Strength, Psychometric Properties, Rasch Model, Response Scaling

Abstract: The summed rating method is popularly known as the Likert scaling model. The characteristics of the data on the Likert scale are the characteristics of ordinal data, or rank order data. One of the efforts to make the data into interval data on the results of psychological measurements is to do a scaling process. This study aims to scaling ordinal data on the Likert scale to logit data using the Rasch model as well as testing the psychometric properties of the Character Strength instrument. Character Strength Measurement, in this study using the Values in Action Inventory Strength of Youth (VIA-Youth) developed by Park and Peterson, consisting of 24 character strengths of 34 items. Respondents in this study were students at Widya Dharma Klaten University, totaling 375 people. The psychometric properties used in this study consisted of item analysis and instrument analysis, using Winstep Software. Item analysis includes: item difficulty level, item suitability level. Instrument analysis includes: unidimensionality, rating scale testing, reliability (respondents and items), and information functions. Overall the psychometric property results of the Character Strength instrument are included in the good category according to the criteria in the Rasch modeling.

1 INTRODUCTION

Scaling is a psychological procedure for placing attributes or characteristics of objects at certain points along a continuum and is expressed quantitatively. By knowing the location or position of an attribute on the continuum quantitatively, the distance between one attribute and another will be known, so that in the end it will also know the quantitative distance of individual differences in the attribute in question. There are three main approaches to scaling, namely an approach with subject-oriented methods, an approach with stimulus-oriented methods, and a response-oriented approach (Torgerson, 1958). Subject-oriented approach to scaling, attention is paid to the problem of placing individuals at points along a continuum. This is generally the purpose of holding aptitude tests and achievement tests. A stimulus-oriented scaling approach aims to place the stimulus on the psychological continuum of the attribute in question. This approach originated from the interest of psychophysicists who tried to find the relationship between the response to a stimulus and the stimulus itself. The most complex scaling approach is the response-oriented approach. Response data is used to place the subject on a psychological continuum based on the strength of the item that is selected or answered correctly (gets endorsement), while at the same time the items are scaled according to the strength or number of properties possessed by the subject who chose or answered correctly earlier.

Response scaling is often used in questionnaires and is the most widely used scale in research in the form of surveys. In response scaling, respondents were asked to complete a questionnaire that required them to indicate their level of agreement with a series of questions. Response scaling is usually used as a measuring tool in educational contexts. Data characteristics on the Likert scale are ordinal data characteristics, or rank order data, so they must be analyzed using non-parametric methods, which are less sensitive and less powerful than parametric ones (Jamieson, 2004). One of the efforts to make the data into interval data on the results of psychological measurements is to do a scaling process.

Suryani, Y. E., Setyowati, F. A., Tukiyo, & Saputri, A. E.
Psychometric Properties of Character Strength Instruments Using the Rasch Model
In *Proceeding of the International Conference on Current Advancement in Psychology (ICCAP) 2022 - Psychology for Sustainable Recovery in the Life after the Pandemic,* page 408-417

The measurement process in quantitative research is very essential because the numbers obtained are used to be processed and analyzed in order to get answers for the research question. Without measurement no data will be obtained and research will not continue. According to the Rasch modeling, a study is related to three things, namely: instruments, items and respondents. In social sciences such as psychology and education, the object being measured is usually an invisible object or something hidden (latent). In Rasch modeling, the basic idea of the measurement process is called latent traits which are none other than the main concept of item response theory (IRT). Even though these characteristics cannot be observed empirically, the manifestation of these characteristics and their interaction with the environment will give rise to an empirical indicator that can be measured or observed. That's why measurement instruments are very important to measure these latent traits.

In the concept of developing psychological measurement instruments, constructs (attributes) related to humans have many characteristics, one of which is a latent construct. That is, these attributes are hypothetical and cannot be observed directly. Methodologically, the use of measurement instruments is a very important part of quantitative research. Reliable and valid instruments will provide reliable information. Conversely, an instrument that fails to meet the requirements will produce biased or misleading results that can reduce the quality of the research. In the social sciences, especially quantitative research, measurement is very important because it is the core of the research activity itself. There are five conditions that must be met to make the test on an instrument truly valid, namely 1) measurable quantity units; 2) a scalable concept; 3) have linear intervals; 4) replicable; and 5) can make predictions (Wright & Mok, 2004). Only Rasch modeling can meet the test requirements of this comprehensive instrument. Therefore, the psychometric property analysis of the Character Strength instrument in this study used Rasch modeling with the help of the Winstep program.

Character Strength Relates to virtue (virtue) possessed by a person as a positive trait (Peterson & Seligman, 2004). Virtue is human goodness, namely virtue in a person that is universal and is reflected through thoughts, feelings, and behaviors that are formed through an evolutionary process and are very important for survival. Character Strength is a specific psychological process that includes broader virtues and is a core characteristic (Niemiec et al., 2017). Character Strength consists of 6 dimensions consisting of 24 character strengths. The first dimension is Intellectual strength, which is a person's cognitive ability towards an expertise, proficiency in use of knowledge that becomes the foundation in the process of achieving a good life. There are four Character Strength indicators on the intellectual strengths dimension, namely (a) curiosity (interest, openness to experience), love of learning, appreciation of beauty and excellence, and creativity. Curiosity is the character of someone who likes all experiences, looking for very interesting topics and subjects, digging to find many things. Love of learning, is an individual character who masters new skills, masters both formal and non-formal knowledge topics. Appreciation of beauty and excellence is an individual character who realizes and appreciates beauty, perfection and performance in all aspects of life, from nature, art, science and mathematics to everyday experiences.

The second dimension of Character Strength is leadership strengths, namely the ability to treat, influence, direct, and motivate others to achieve success. Indicators on the dimensions of leadership strengths consist of leadership, humor, perspective, social intelligence, and bravery. Leadership is a character that encourages people in a group to work while maintaining good relations with group members. Humor is a character possessed by individuals who like to make other people smile, make positive jokes. Perspective, namely the ability of individuals to give advice, have a broad perspective and can be accepted by others. Social intelligence is an individual character that is sensitive to the motives and feelings of other people and oneself, can adapt to different situations, and knows how to move others. Bravery is an individual character who is not afraid of threats, challenges or pain, dares to express wishes even though there are opponents, dares to be different even though he is not popular.

The third dimension of Character Strength is Other-directed strengths, namely positive traits which are the individual's ability to maintain interpersonal relationships and the ability to pay attention to one's rights and obligations in group life. There are five indicators on the other-directed strengths dimension, namely modesty (humanity), forgiveness and mercy, kindness, fairness, and teamwork (loyalty). Modesty (humanity) is a character possessed by individuals to achieve, not seek attention, and do not consider themselves more special than other people. Forgiveness and mercy character to forgive others who do wrong, give others a chance, not hold grudges. Kindness, namely

doing good to others, helping others, and taking care of others. Fairness, namely treating everyone fairly, influencing decisions that concern other people, giving equal opportunity to everyone. Team work (loyalty) individual character who can work well in a group, loyal to the group, and share with the group.

The fourth dimension is Temperance strengths, namely the ability to hold back and not do something that is considered excessive. Indicators on the temperance strengths dimension, namely prudence, self-regulation, prevalence, open-mindedness, and honesty (integrity). Prudence is a prudent character, individuals are careful with the decisions they make, don't take unnecessary risks, don't do anything irresponsible. Self-regulation, namely regulating feelings and behavior, discipline, controlling emotions. Preverence, namely completing the work that has been started, diligently carrying out activities even though there are obstacles, focusing on the goals to be achieved, happy in completing tasks. Open-mindedness, namely thinking from all perspectives, being careful in drawing conclusions, considering all evidence/possibilities fairly, being able to change one's mind to concrete evidence. Honesty (integrity) is an individual who conveys the truth, presents himself as he is, is responsible for his feelings and behavior.

The fifth dimension is Transcendence strengths, namely the individual's ability to establish a relationship with the wider universe's forces in making sense of life. There are five indicators on the transcendence strengths dimension, namely religiousness (spirituality), zest (vitality, enthusiasm), gratitude, love, hope (optimism). Religiousness (spirituality), namely individuals who have coherent beliefs about a higher will and the meaning of the universe, have beliefs about the meaning of life that shape behavior and provide comfort. live life as if adventurous, full of enthusiasm / active. Gratitude, individuals who are aware and grateful for the good things that happen, take the time to express gratitude. love, namely the character of respecting relationships with others, sharing and caring, close to others. Hope (optimism), namely expecting the best for the future and trying to make it happen, realizing that fate can change and a good future can be achieved.

The last (sixth) dimension in Character Strength is Metacognitive strengths, namely the ability to think in order to build strategies to solve problems as a basis for deciding problems. There are two indicators of Character Strength on the metacognitive strengths dimension, namely strategy and goals. Strategy is an individual who has an efficient and effective learning style, knows your weaknesses and is able to overcome them. Goal, namely the individual character that is able to organize activities to achieve goals, and know the goals to be achieved (Peterson & Seligman, 2004).

2 METHODS

The type of research used is descriptive to describe the characteristics of the Character Strength instrument. The Character Strength instrument has 6 dimensions which reflect 24 character strengths, namely (1) intellectual strengths consisting of the characters of curiosity, love of learning, appreciation of beauty and excellence, and creativity); (2) leadership strengths (leadership, humor, perspective, social intelligence, and courage); (3) other-directed strengths (modesty, forgiveness, kindness, fairness, and teamwork); (4) temperance strengths (prudence, self-regulation, perseverance, open-mindedness, and honesty); and (5) transcendence strengths (religiousness, zest, gratitude, love, and hope, (6) metacognitive strengths (goals and strategic) (Park & Peterson, 2009). This measuring instrument character strength itself uses a response scaling with five responses, namely Very Appropriate, Appropriate, Fairly Appropriate, Not Appropriate, Very Unsuitable. Respondents in this study were 375 students at Widya Dharma Klaten University.

The data generated from the Character Strength measurement is ordinal data, the scaling process is carried out so that it can be used in parametric statistical analysis. Scaling is done to transform ordinal data into interval data, in this study the data is transformed to a logit scale using the Rasch model. The calibration process uses the help of Winstep software (Linacre, 2011). The use of the Rasch model for polytomous data was developed by Andrich while still being based on two basic theorems, namely the level of individual ability/agreement and the difficulty level of the item to agree on (Misbach & Sumintono, 2014). The psychometric properties used in this study include item analysis and instrument analysis. Item analysis consists of item difficulty level, item suitability level, and detection of item bias. The criteria used to see item fit are Outfit means square, Outfit Z-standard, and Point measure correlation. If these three criteria are not met, then the item is said to be misfit (not suitable for the Rasch model). The criteria for item fit are as follows:

Outfit Mean Square Residual (MNSQ): 5<y<1.5

Outfit Standardized Mean Square Residual (ZSTD): -2 < Z < + 2; and
Point Measure Correlation (PTMEA CORR): 0.4 < x < 0.8 (Boone et al., 2014)

Instrument analysis in Rasch modeling consists of: unidimensionality, rating scale testing, reliability (respondents and items), and information functions.

3 RESULTS

3.1 Item Analysis Results

The Character Strength instrument is considered to have good Character Strength, if it is able to measure Character Strength. In item analysis using the Rasch model, there are 2 things that are produced, namely the difficulty level of the items (item measure), the level of suitability of the items (item fit), and item bias. The results of the item difficulty level analysis can be seen from the measure values obtained. Winstep's output will sort the items from the highest to the lowest logit value. A high logit value indicates a high item difficulty level. Based on the results of data analysis, it can be seen that the item that has the highest level of difficulty is item 30 with a logit value of 1.31 and the easiest item is item 2 with a logit value of -2.24. Other information obtained by Rasch modeling is looking at the quality of the fit of the items with the model or what is called item fit. Item fit explains whether the item is functioning normally when measuring or not. The MNQS logit value shows how big the misfit impact of the item is, which is seen in two forms, namely Outfit MNSQ and Infit MNSQ Outfit is chi-squared which is sensitive to outliers (deviations) in the Rasch model analysis. Outliers are often a lucky guess for lower ability respondents and a careless mistake for higher ability students. The mean squares that are inappropriate or misfit tend to be influenced by response patterns, which focus on responses that approach the difficulty of the item or the ability of the respondent (Lu & Bi, 2016). Expected MNSQ value is 1.0 logit. The value of PT MEA CORR is the correlation between item person score and person size. For the Rasch model, the PT MEA CORR value must be positive and not close to zero (Bond, Yan, & Heenee, 2020). Based on table 1 above, it can be seen that: (a) item 1 does not match the MNSQ Outfit criteria; (b) items 1, 2, 3, 9, 11, 16, 17, 18, 20, 21, 23 and 24 are not appropriate The value of PT MEA CORR is the correlation between item person score and person size. For the Rasch model, the PT MEA CORR value must be positive and not close to zero (Bond, Yan, & Heenee, 2020). There is 1 item that does not match the MNSQ Outfit criteria. Items 1, 2, 3, 9, 11, 16, 17, 18, 20, 21, 23 and 24 are not appropriate with ZSTD Outfit criteria. Items 1, 2, 6, 14, 26 do not meet the PTMEA CORR criteria. Thus there is 1 item that does not meet the three criteria, namely item 1, so it can be concluded that item 1 does not fit the Rasch model (misfit). These results indicate that the Character Strength instrument has good measurement effectiveness. In addition to using lift, item discrepancies in Rasch modeling can be displayed via a graph as shown below.

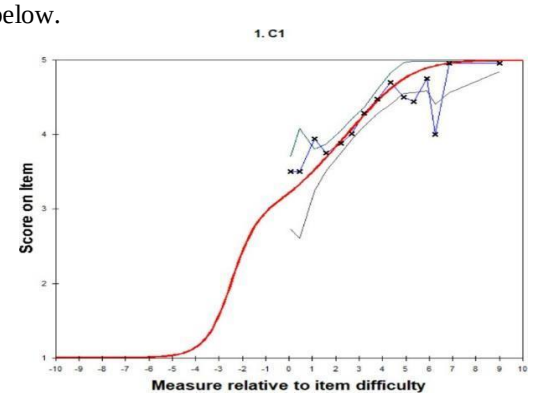


Figure 1. Misfit Items.

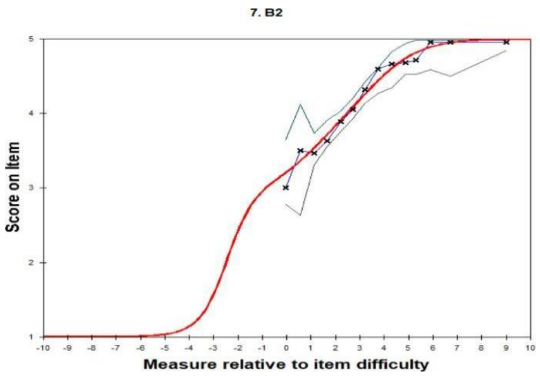


Figure 2. Item Fit.

Based on the figure it can be seen that there are responses that come out of the upper black line (trust space of the outfit curve) and come out of the bottom black line (trust space of the infit curve). This indicates that the item is misfit (item 1). For

item fit, there will be no response outside the infit or outfit confidence curve, as in item 7, as shown in Figure 6 below.

An instrument or item is said to be biased if it is found that one individual with certain characteristics is more advantageous than an individual with other characteristics. In Rasch modeling to detect this bias is called detection DIF (differential item functioning or differential item functioning). The item bias in this measurement is seen based on one variable, namely gender. The Rasch model analysis features the detection of item bias in Differential Item Functioning (DIF). Bias can be identified based on the probability value of items that are below 5% (Sumintono & Widhiarso, 2015a). Referring to the results of the DIF analysis based on gender, it was found that there was one item that contained a bias, namely item 2. This meant that item 2 was perceived differently by students and female students. While the other 33 items can be believed that the items are perceived the same by respondents of different sexes.

3.2 Instrument Analysis Results

In Rasch modeling, besides being able to carry out analysis down to the item level and individual level, analysis at the instrument level can also be carried out. Another advantage, the analysis of the instrument is given in more detail in the form of statistical summaries and function test information. This means that the value of the information provided by Rasch modeling is comprehensive. Rasch modeling provides two estimates of instrument reliability, namely person and item reliability. Reliability testing aims to determine how far the measurement produces consistent information in uncovering latent traits or the nature of the unidimensional variable being measured (Sumintono & Widhiarso, 2015).

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	135.4	34.0	2.44	.31	1.01	-.17	1.01	-.17
SEM	.6	.0	.06	.00	.02	.10	.02	.10
P.SD	12.2	.0	1.14	.03	.44	1.94	.46	1.89
S.SD	12.2	.0	1.14	.03	.44	1.94	.46	1.89
MAX.	167.0	34.0	6.35	.61	3.69	7.81	3.65	6.65
MIN.	104.0	34.0	-.55	.29	.20	-5.44	.24	-5.07

REAL RMSE	.33	TRUE SD	1.09	SEPARATION	3.27	Person RELIABILITY	.91
MODEL RMSE	.31	TRUE SD	1.10	SEPARATION	3.56	Person RELIABILITY	.93
S.E. OF Person MEAN = .06							

Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .93 SEM = 3.34
STANDARDIZED (50 ITEM) RELIABILITY = .95
SUMMARY OF 34 MEASURED (NON-EXTREME) Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	1494.4	375.0	.00	.09	1.00	-.10	1.01	-.03
SEM	15.9	.0	.14	.00	.03	.39	.03	.39
P.SD	91.3	.0	.79	.00	.16	2.25	.18	2.26
S.SD	92.7	.0	.80	.00	.16	2.28	.18	2.30
MAX.	1739.0	375.0	1.31	.11	1.48	6.22	1.55	6.80
MIN.	1338.0	375.0	-2.24	.09	.74	-4.19	.73	-4.26

REAL RMSE	.10	TRUE SD	.78	SEPARATION	8.20	Item RELIABILITY	.99
MODEL RMSE	.09	TRUE SD	.78	SEPARATION	8.46	Item RELIABILITY	.99
S.E. OF Item MEAN = .14							

Item RAW SCORE-TO-MEASURE CORRELATION = -1.00

Figure 3. Reliability Analysis Results.

Based on Figure 4 it can be seen that the Cronbach Alpha reliability coefficient value obtained is 0.93. This means that the instrument is considered to have a high level of consistency so that it can be used in actual research because it meets the minimum requirements of >0.7 (Sumintono & Widhiarso, 2015b). Acquiring a Cronbach Alpha Coefficient (KR-20) value of 0.93 indicates that there is good instruction between students and the Character Strength instrument. This means that there is a very strong correlation between student responses and items, where student knowledge tends not to be fragmented so that it can be measured (Adams & Wieman, 2011). It can be said that the Character Strength instrument is sensitive and reliable enough to distinguish well the Student's Character Strength. The person reliability value was 0.98, while the item reliability value was 0.99 which was also included in the high category. This shows that the consistency of the items is very good, or items can be said to meet the unidimensionality requirements. This means that the item is able to define the variable being measured very well. This conclusion is confirmed through the acquisition of infit and outfit item values, most of which are in the acceptable range (Bond, Yan, & Heenee, 2020; Herrmann-abell & Deboer, 2016). The average measure value obtained in the person table is 2.44 to fill in the Character Strength scale. The average value that is greater than the logit value of 0.0 indicates a tendency for student abilities to be greater than the difficulty level of the instrument. The average Infit MNQS and Outfit MNQS person obtained is 1.01 points 98 while the item reliability value of 0.99 is also included in the high category. This shows that the consistency of the items is very good, or items can be said to meet the unidimensionality requirements. This means that the item is able to define the variable being measured very well. This conclusion is confirmed through the acquisition of infit and outfit item values, most of which are in the acceptable range (Bond, Yan, & Heenee, 2020; Herrmann-abell & Deboer, 2016). The average measure value obtained in the person table is 2.44 to fill in the Character Strength scale. The average value that is greater than the logit value of 0.0 indicates a tendency for student abilities to be greater than the difficulty level of the instrument. The average Infit MNQS and Outfit MNQS person obtained is 1.01 points 98 while the item reliability value of 0.99 is also included in the high category. This shows that the consistency of the items is very good, or items can be said to meet the unidimensionality requirements. This means that the

item is able to define the variable being measured very well. This conclusion is confirmed through the acquisition of infit and outfit item values, most of which are in the acceptable range (Bond, Yan, & Heenee, 2020; Herrmann-abell & Deboer, 2016). The average measure value obtained in the person table is 2.44 to fill in the Character Strength scale. The average value that is greater than the logit value of 0.0 indicates a tendency for student abilities to be greater than the difficulty level of the instrument. The average Infit MNQS and Outfit MNQS person obtained is 1.01 points 99 is also included in the high category. This shows that the consistency of the items is very good, or items can be said to meet the unidimensionality requirements. This means that the item is able to define the variable being measured very well. This conclusion is confirmed through the acquisition of infit and outfit item values, most of which are in the acceptable range (Bond, Yan, & Heenee, 2020; Herrmann-abell & Deboer, 2016). The average measure value obtained in the person table is 2.44 to fill in the Character Strength scale. The average value that is greater than the logit value of 0.0 indicates a tendency for student abilities to be greater than the difficulty level of the instrument. The average Infit MNQS and Outfit MNQS person obtained is 1.01 points 99 is also included in the high category. This shows that the consistency of the items is very good, or items can be said to meet the unidimensionality requirements. This means that the item is able to define the variable being measured very well. This conclusion is confirmed through the acquisition of infit and outfit item values, most of which are in the acceptable range (Bond, Yan, & Heenee, 2020; Herrmann-abell & Deboer, 2016). The average measure value obtained in the person table is 2.44 to fill in the Character Strength scale. The average value that is greater than the logit value of 0.0 indicates a tendency for student abilities to be greater than the difficulty level of the instrument. The average Infit MNQS and Outfit MNQS person obtained is 1.01 points or otem can be said to meet the requirements of unidimensionality. This means that the item is able to define the variable being measured very well. This conclusion is confirmed through the acquisition of infit and outfit item values, most of which are in the acceptable range (Bond, Yan, & Heenee, 2020; Herrmann-abell & Deboer, 2016). The average measure value obtained in the person table is 2.44 to fill in the Character Strength scale. The average value that is greater than the logit value of 0.0 indicates a tendency for student abilities to be greater than the difficulty level of the instrument. The average Infit MNQS and Outfit MNQS person obtained is 1.01 points or otem can be said to meet the requirements of unidimensionality. This means that the item is able to define the variable being measured very well. This conclusion is confirmed through the acquisition of infit and outfit item values, most of which are in the acceptable range (Bond, Yan, & Heenee, 2020; Herrmann-abell & Deboer, 2016). The average measure value obtained in the person table is 2.44 to fill in the Character Strength scale. The average value that is greater than the logit value of 0.0 indicates a tendency for student abilities to be greater than the difficulty level of the instrument. The average Infit MNQS and Outfit MNQS person obtained is 1.01 points 0 indicates a tendency for student abilities to be greater than the difficulty level of the instrument. The average Infit MNQS and Outfit MNQS person obtained is 1.01 points 0 indicates a tendency for student abilities to be greater than the difficulty level of the instrument. The average Infit MNQS and Outfit MNQS person obtained is 1.01 points; this exceeds the ideal value of 1.00 because the closer to 0 the better (Sumintono & Widhiarso, 2015b). The average Infit ZSTD and Outfit ZSTD obtained is -0.17 which is an ideal value of 0.0 (the closer to 0.0 the better). Thus it can be concluded that the respondents who answered the scale, in this case the Widya Dharma Klaten University students, answered the scale seriously and the quality of their answers was good.

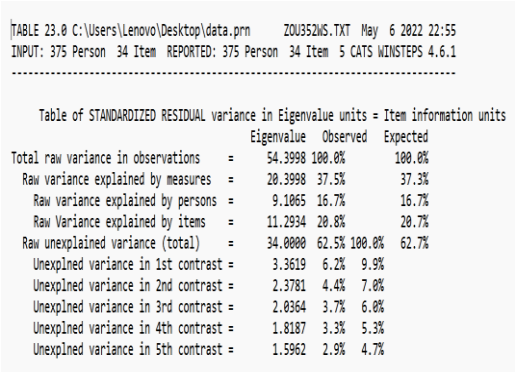
Person separation index estimates how well the Character Strength instrument can differentiate student opinions. The greater the value of the person separation index indicates how widely the items are spread in defining difficult and easy items. If the separation index exceeds 2, it can be assumed to have a good value (Linacre, 2012). The wider the distribution of items, the better and more appropriate. In this study, the person separation index (3.27 logit), item separation (8.20 logit), both reflect the distribution of the Character Strength instrument which is quite good for students and items. This criterion supports the character strength instrument as an appropriate and reliable instrument for knowing student Character Strength.

$$H = \frac{[(4\ x\ SEPARATION) + 1]}{3} = \frac{[(4\ x\ 3{,}27) + 1]}{3} = 4{,}6$$

Based on this formula, the value of H = 4.6 is obtained. This shows that respondents can be divided into four major groups, namely groups that have very high, high, low, and very low Character Strength values. The same method was also used to calculate the separation item index, obtaining a

separation item index value of 11.26. This means that the items in the Character Strength instrument can be grouped into 11 item difficulty levels. This can be interpreted that the items in the Character Strength instrument have good variants.

Measurement dimensionality testing is useful for optimizing the measurements taken, so that the resulting information is more focused on the attributes being measured. The unidimensionality of the instrument is an important measure for evaluating whether the developed instrument is capable of measuring what it should measure. Rasch modeling uses principal component analysis (Principal Component Analysis) of the residuals, which measures the extent to which the diversity of the instrument measures what should be measured (Sumintono & Widhiarso, 2015a). The Character Strength instrument used in this study, relatively, has a good measure of unidimensionality. The results of the analysis are presented in the figure below.



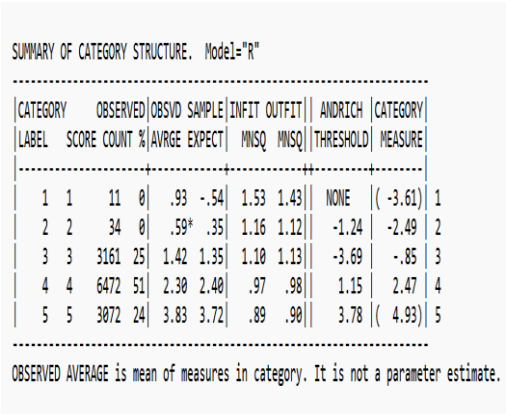
```
TABLE 23.0 C:\Users\Lenovo\Desktop\data.prn      ZOU352WS.TXT  May  6 2022 22:55
INPUT: 375 Person  34 Item  REPORTED: 375 Person  34 Item  5 CATS WINSTEPS 4.6.1
-------------------------------------------------------------------------------

    Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units
                                    Eigenvalue   Observed   Expected
Total raw variance in observations   =    54.3998 100.0%        100.0%
  Raw variance explained by measures =    20.3998  37.5%         37.3%
    Raw variance explained by persons =    9.1065  16.7%         16.7%
    Raw Variance explained by items   =    11.2934  20.8%         20.7%
  Raw unexplained variance (total)    =    34.0000  62.5% 100.0%   62.7%
    Unexplned variance in 1st contrast =    3.3619   6.2%   9.9%
    Unexplned variance in 2nd contrast =    2.3781   4.4%   7.0%
    Unexplned variance in 3rd contrast =    2.0364   3.7%   6.0%
    Unexplned variance in 4th contrast =    1.8187   3.3%   5.3%
    Unexplned variance in 5th contrast =    1.5962   2.9%   4.7%
```

Figure 4. Unidimensionality.

In the picture above it can be seen that the raw variance data measurement results are 37.5%. This value indicates that the unidimensionality requirement is very good. The minimum value that must be met for this requirement is 20%, and if it reaches more than 60% it means special. Other values of raw unexplained (variant that cannot be explained by the instrument), ideally not exceeding 15%. The results obtained are below 10%. This means that the Character Strength instrument can effectively measure student Character Strength.

The measurement of the rating scale aims to test whether the assessment criteria or ranking of student responses can be used or not. Rasch modeling can provide a verification process for the rating assumptions used in the Character Strength instrument. Rating size is seen from the observed average and Andrich Threshold. The results of the analysis can be seen in the image below.



SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	OBSERVED SCORE	COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	ANDRICH THRESHOLD	CATEGORY MEASURE	
1	1	11	0	.93	-.54	1.53	1.43	NONE	(-3.61)	1
2	2	34	0	.59*	.35	1.16	1.12	-1.24	-2.49	2
3	3	3161	25	1.42	1.35	1.10	1.13	-3.69	-.85	3
4	4	6472	51	2.30	2.40	.97	.98	1.15	2.47	4
5	5	3072	24	3.83	3.72	.89	.90	3.78	(4.93)	5

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

Figure 5. Rating Scale Analysis.

In the picture above it can be seen that the observed Average starts from logit 0.93 for rank 1 (Strongly Agree), then decreases to logit 0.59 for rank 2 (Agree), then increases again for ratings 3, 4 and 5. From rank 1 to 5. The transition to the Character Strength ranking category occurs consistently (Linacre, 2020). This shows that the 1-5 rating scale is not confusing for respondents and is the right scaling range in this instrument. Another recommended measure is the Andrich Threshold to test whether the polytomous value used is correct or not. The Andrich Threshold value which moves from None to negative and leads to positive sequentially indicates that the five options given are valid for the respondent.

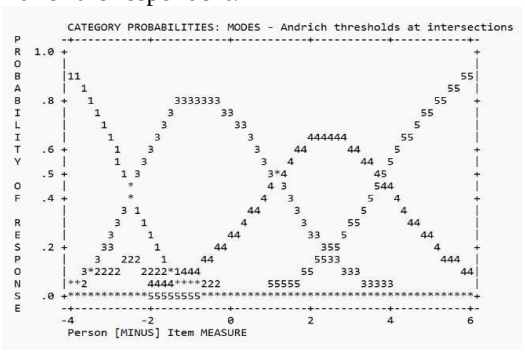


Figure 6. Rating Scale Analysis Graph.

Instrument Character Strength can be well understood by respondents because according to the peak criteria of each scale (rating), it can be seen that each rating (1, 2. 3. 4. 5) has a separate peak. Paying attention to these data, it can be stated that the probability of each rating is clearly different from the respondents. This means that the Character Strength instrument being tested was able to differentiate the ratings by the respondents well.

Each measurement always produces information about the measurement results. This measurement information depends on the relationship between the test and the individual being measured. Measurement information is affected by the variation in the results obtained. The information function in the item response theory that fits the model can be divided into two, namely: item information function for items and test information function for tests. In classical test theory, item analysis is used to select adequate items for the group of test takers, whereas in item response theory to select adequate items is by using the item information function. In classical test theory the reliability coefficient is used to assess the quality of a test set consisting of a number of items, whereas in item response theory to achieve the same goal is obtained through the information function of the test. The information function is used to describe the strength of items or tests, selection of test items and comparing test sets (Hambleton et al., 1991). The information function states the strength or contribution of the test in uncovering the latent trait measured by the test. The results of the analysis of the measurement information function on the Character Strength instrument can be seen in the graph below.

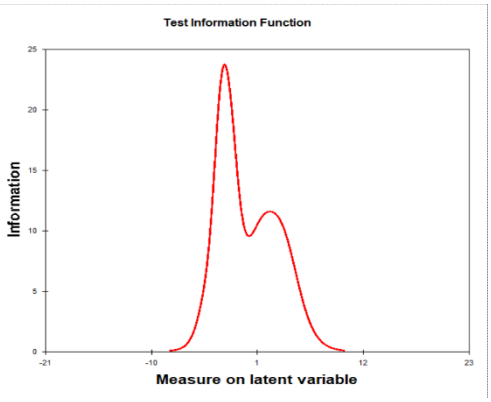


Figure 7. Information Function Graph.

From the graph of the information function above, it can be seen that there are two peaks of optimal information obtained by the test, namely individuals with very high Character Strength and individuals with moderate Character Strength. Based on the information function graph, it can also be shown that the reliability of the Character Strength instrument is very high. The higher the peak of the information function, the higher the value of the reliability of the measurements carried out (Su Mintano & Widhiarso, 2015).

4 DISCUSSION

Measurement is a very fundamental activity in science. Knowledge about people, objects, events and processes is obtained through observation. Reasonable interpretation of all observations must be made through measurement. Without measurement there will be no data obtained. The measurement process is the process of entering numbers into the concept according to the indicators that have been made when the continuum has been determined (Azwar, 2020). In social sciences such as Psychology and Education, the object being measured is usually an invisible object (visible) or something hidden (latent). Variables or attributes in measurement are psychological attributes in the form of ability attributes and non-ability attributes. All attributes as objects in measurement are theoretical,*Character Strength* developed as an effort to shape one's quality of life so that it can be better (Peterson & Seligman, 2004). Like personality, a person's Character Strength tends to be stable over time (Gander et al., 2020). Character Strength is the foundation for optimal self-development during one's life and are also aspects of personality that are morally valued (Park & Peterson, 2009). The 24 character strengths can be applied in various aspects of life. Character Strength that is applied in the workplace has a positive correlation with job satisfaction, productivity, organizational citizenship behavior, in a person (Heintz et al., 2021; Lavy & Littman-Ovadia, 2016). Character Strength is also positively correlated with the five indicators of happiness, namely positive affect, engagement, quality relationships, meaningfulness, and achievement (Wagner et al., 2020). The use of appropriate character strengths can increase one's life satisfaction (Douglass & Duffy, 2015). Character Strength is an important construct in psychology, especially positive psychology, which will be very beneficial for both individual and organizational development. The Character Strength instrument is a measuring tool that can be used to study positive qualities in humans (Garvin, 2020). Many studies have been conducted on the quality of the Character Strength instrument.

Psychological attributes such as intelligence, talent, academic achievement, and Character Strength cannot be measured (quantified) as directly as physical attributes. Psychological attributes can only be revealed through various forms of behavior that indicate the presence of these attributes in a person. Behavioral indicators must be well formulated so that they can be used as a reference in

developing instruments so that the existence of psychological attributes can be quantified using scales or psychological tests. In Rasch modeling, the basics of the measurement process like this are called latent traits which are the main concepts of item response theory (IRT). Although these characteristics cannot be observed empirically, the manifestation of these characteristics and their interaction with the environment will give rise to an empirical indicator that can be measured or observed, the instruments developed or used for measurement can infer the presence or absence of these characteristics. Therefore the instrument is made using various items (items or item) as a basis for making conclusions about latent traits. The instruments used must produce accurate data so that the decisions made are accurate (Mardapi, 2017). For this reason, the instruments used must be of high quality. The quality of the instrument is seen from two main things, namely validity and reliability. There are many methods that can be used to test the quality of the instrument, you can use classical test theory or modern test theory.

Based on the results of the analysis of the psychometric properties of the strength instrument, it can be concluded that the Character Strength instrument has good quality, both the items and the instrument. Testing of the factors in the character strength instrument has also been carried out in various countries (Ng et al., 2017; Ruch et al., 2014; van Eeden et al., 2008; Dametto & Noronha, 2019). Validity and reliability are classic test theory concepts which are still widely used today. The quality of the Character Strength instrument using classical theory states that the Character Strength instrument has good quality (Garvin, 2020). Testing the quality of the Character Strength instrument in this research uses the Rasch modeling, the results show that the quality of this character strength instrument is of good quality, these results are in line with previous research (Hanim & Herdi, 2022). In testing the quality of modeling instruments Rasch has advantages over other methods, especially classical test theory because of its ability to predict missing data, which is based on systematic response patterns (Wu & Adams, 2007; Kyngdon, 2008; Strobl, Kopf, & Zeileis, 2015). Rasch modeling is capable of producing standard error measurement values for the instruments used so that they can be used to increase the accuracy of calculations (Bond, Yan, & Heenee, 2020). In Rasch modeling, calibration is carried out simultaneously on the measurement scale, responses and item items. Instruments that have been calibrated will produce valid data, instruments can measure what is to be measured so that the results of measurements or research carried out can be trusted.

5 CONCLUSION

Overall the results of the psychometric properties of the Character Strength instrument, which consist of: unidimensionality, rating scale testing, reliability (respondents and items), and information functions are included in the good category and in accordance with the criteria in Rasch modeling. This study only used a limited sample, so that future research is expected to use a larger number of samples with more diverse demographics, in terms of age, culture, place of residence and so on.

REFERENCE

Adams, W. K., & Wieman, C. E. (2011). Development and Validation of Instruments to Measure Learning of Expert-like Thinking. *International Journal of Scince Education*, *33*(9), 1289–1312. https://doi.org/10.1080/09500693.2010.512369

Azwar, S. (2020). *Dasar-Dasar Psikometrika* (2nd ed.). Pustaka Pelajar.

Bond, T. ., Yan, Z., & Heenee, M. (2020). *Applying The Rasch Model: Fundamental Measurement in the Human Sciences* (4th ed.). Routledge. https://doi.org/10.4324/9780429030499

Dametto, D. M., & Noronha, A. P. P. (2019). Construction and Validation of the Character Strengths Scale for Youth (CSS-Youth). *Paideia*, *29*(2000), 1–10. https://doi.org/10.1590/1982-4327E2930

Douglass, R. P., & Duffy, R. D. (2015). Strengths Use and Life Satisfaction: A Moderated Mediation Approach. *Journal of Happiness Studies*, *16*(3), 619-632. https://doi.org/10.1007/s10902-014-9525-4

Gander, F., Hofmann, J., Proyer, R. T., & Ruch, W. (2020). Character Strengths – Stability, Change, and Relationships with Well-Being Changes. *Applied Research in Quality of Life*, *15*(2), 349–367. https://doi.org/10.1007/s11482-018-9690-4

Garvin, G. (2020). Penyusunan Alat Ukur Character Strength Bahasa Indonesia. *PSYCHE: Jurnal Psikologi*, *2*(2), 28–48. https://doi.org/10.36269/psyche.v2i2.238

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. SAGE Publications. https://books.google.co.id/books?id=cmJU9SHCzecC

Hanim, W., & Herdi, H. (2022). Adaptation and Validation of The Character Strength Inventory for Indonesian Students. *Jurnal Pendidikan dan Pengajaran*, *55*(1), 204-210.

Heintz, S., Ruch, W., Lau, C., Saklofske, D. H., & McGhee, P. (2021). Development and Validation of the Short Version of the Sense of Humor Scale (SHS-S). *European Journal of Psychological Assessment*. https://doi.org/10.1027/1015-5759/a000670

Herrmann-abell, C. F., & Deboer, G. E. (2016). *Herrmann-Abell & DeBoer, AERA 2016*. 1–12.

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, *38*(12), 1217–1218. https://doi.org/10.1111/j.1365-2929.2004.02012.x

Kyngdon, A. (2008). The Rasch Model from the Perspective of the Representational Theory of Measurement. *Theory & Psychology*, *18*(1), 89–109. https://doi.org/10.1177/0959354307086924

Lavy, S., & Littman-Ovadia, H. (2016). My Better Self: Using Strengths at Work and Work Productivity, Organizational Citizenship Behavior, and Satisfaction. *Journal of Career Development*, *44*(2), 95–109. https://doi.org/10.1177/0894845316634056

Linacre, J. M. (2011). *Winsteps Help for Rasch Analysis*. http://homes.jcu.edu.au/~edtgb/%5Cnpapers3://publication/uuid/D56B724A-62FF-4D00-84E1-ECC888298B70

Linacre, M. (2012). Winsteps Tutorial 3. *Winstep Tutorial*, *June*, 1–38. https://www.winsteps.com/a/winsteps-tutorial-1.pdf

Linacre, J. M. (2020). T*he transition to the Character Strength ranking category occurs consistently*. [Psychometric assessment of character strengths]. In *Psychological measurement and character strengths*.

Lu, S., & Bi, H. (2016). Development of a measurement instrument to assess students' electrolyte conceptual understanding. *Chemistry Education Research and Practice*, *17*(4), 1030–1040. https://doi.org/10.1039/C6RP00137H

Mardapi, D. (2017). *Pengukuran, Penilaian dan Evaluasi Pendidikan* (2nd ed.). Parama Publishing.

Misbach, I. H., & Sumintono, B. (2014). Pengembangan dan validasi instrumen "persepsi siswa tehadap karakter moral guru" di Indonesia dengan model rasch. *PROCEEDING Seminar Nasional Psikometri*, *May*, 148–162.

Ng, V., Cao, M., Marsh, H. W., Tay, L., & Seligman, M. E. (2017). The Factor Structure of the Values in Action Inventory of Strengths (VIA-IS): An Item-Level Exploratory Structural Equation Modeling (ESEM) Bifactor Analysis. *Psychological assessment*, *29*(8), 1053.https://doi.org/10.1037/pas0000396

Niemiec, R. M., Shogren, K. A., & Wehmeyer, M. L. (2017). Character strengths and intellectual and developmental disability: A strengths-based approach from positive psychology. *Education and Training in Autism and Developmental Disabilities*, *52*(1), 13–25.

Park, N., & Peterson, C. (2009). Character Strengths: Research and Practice. *Journal of College & Character*, *10*(4), 1–10.

Peterson, C., & Seligman, M. E. . (2004). *Chacter Strengths and Virtues*. Oxford University Press.

Ruch, W., Weber, M., Park, N., & Peterson, C. (2014). Character strengths in children and adolescents: Reliability and initial validity of the German Values in Action Inventory of Strengths for Youth (German VIA-Youth). *European Journal of Psychological Assessment*, *30*(1), 57–64. https://doi.org/10.1027/1015-5759/a000169

Strobl, C., Kopf, J., & Zeileis, A. (2015). Rasch Trees: A New Method for Detecting Differential Item Functioning in the Rasch Model. *Psychometrika*, *80*(2), 289–316. https://doi.org/10.1007/s11336-013-9388-3

Sumintono, B., & Widhiarso, W. (2015a). *Aplikasi Mosel Rasch Untuk Penelitian Ilmu-Ilmu Sosial*. Tri Komunikata.

Sumintono, B., & Widhiarso, W. (2015b). *Aplikasi Pemodelan Rasch Pada Assesment Pendidikan*. Tri Komunikata.

Torgerson, W. S. (1958). Theory and methods of scaling. In *Theory and methods of scaling*. Wiley.

van Eeden, C., Wissing, M. P., Dreyer, J., Park, N., & Peterson, C. (2008). Validation of the Values in Action Inventory of Strengths for Youth (VIA-Youth) Among South African learners. *Journal of Psychology in Africa*, *18*(1), 143–154. https://doi.org/10.1080/14330237.2008.10820181

Wagner, L., Gander, F., Proyer, R. T., & Ruch, W. (2020). Character Strengths and PERMA: Investigating the Relationships of Character Strengths with a Multidimensional Framework of Well-Being. *Applied Research in Quality of Life*, *15*(2), 307–328. https://doi.org/10.1007/s11482-018-9695-z

Wright, B. D., & Mok, M. M. C. (2004). Chapter 1: An Overview of the Family of Rasch Measurement Models. *Introduction to Rasch Measurement: Theory, Models and Applications*, *1*(1) 1–24.

Wu, M., & Adams, R. (2007). *Applying the Rasch model to psycho-social measurement: A practical approach*. Melbourne: Educational Measurement Solution