International Journal of Evaluation and Research in Education (IJERE)
Vol. 13, No. 2, April 2024, pp. 672~681
ISSN: 2252-8822, DOI: 10.11591/ijere.v13i2.26517

Analysis of the Indonesian version of the statistical anxiety scale instrument with a psychometric approach

Deni Iriyadi[1], Ahsanul Khair Asdar[2], Bambang Afriadi[3], Mohammad Ardani Samad[4], Yogi Damai Syaputra[5]
[1]Postgraduate Program, Universitas Islam Negeri Sultan Maulana Hasanuddin Banten, Serang, Indonesia
[2]Postgraduate Program, Sekolah Tinggi Agama Budha Negeri Sriwijaya Tangerang, Tangerang, Indonesia
[3]Department of Management, Faculty of Economy, Universitas Islam Syekh Yusuf, Tangerang, Indonesia
[4]Department of Hospital Administration, Faculty of Health Sciences, Institut Ilmu Kesehatan Pelamonia Kesdam XIV/Hasanuddin, Makassar, Indonesia
[5]Deparment of Islamic Counseling Guidance, Faculty of Da'wah, Universitas Islam Negeri Sultan Maulana Hasanuddin Banten, Serang, Indonesia

Article Info

Article history:

Received Jan 6, 2023
Revised Sep 14, 2023
Accepted Oct 4, 2023

Keywords:

Factor analysis
Rasch model
Rating scale model
Statistical anxiety
Statistical anxiety scale

ABSTRACT

This study aims to validate the psychometric approach (using the Rasch model and confirmatory factor analysis) on the Indonesian version of the statistical anxiety scale (SAS) instrument. Sampling in this study used cluster random sampling with a total sample of 1651 which was divided into three regions of Indonesia (Western Indonesia, Central Indonesia, and Eastern Indonesia) with details of 457 males (27.46%) and 1,194 females (72.54%). The number of samples from the West Indonesia region was 922 people (55.71%), the Central Indonesia region was 605 people (36.76%), and the Eastern Indonesia region was 124 people (7.53%). The results of confirmatory factor analysis (CFA) show that the dimensions/elements of the SAS instrument meet the criteria which are then continued using the rating scale model (RSM) approach. From the results of the tests performed, it shows that all items on the SAS instrument meet the specified criteria. Thus, the SAS instrument adopted in the Indonesian context can be used to measure students' statistical anxiety. The results of the research conducted became the initial capital to increase students' statistical anxiety considering that statistics courses are needed in completing studies, especially quantitative research.

Corresponding Author:

Deni Iriyadi
Postgraduate Program, Universitas Islam Negeri Sultan Maulana Hasanuddin Banten
Jend. Sudirman street No. 30, Serang - Banten 42118, Indonesia
Email: deni.iriyadi@uinbanten.ac.id

1. INTRODUCTION

Statistics is the study of how to collect data, process data, analyze the data that has been collected to draw conclusions [1], [2]. This process is certainly carried out with the aim that the conclusions obtained are in accordance with the facts on the ground without any interference from any party. Statistics is a record of numbers (numbers). Not just numbers, experts agree that statistics are numbers that are collected, tabulated, and classified to be able to provide accurate and precise information [3], [4]. The function of statistics is to solve problems practically. There are two statistical functions to be aware of. Distinct measurable capability, to portray or make sense of information and occasions. Then, at that point, the inferential capability, to anticipate and control the whole populace in view of information, side effects, and occasions contained in the exploration cycle [5]. Measurements in Education, with extraordinary thought for idea estimation and assessment, is a significant piece of the educating and growing experience [6]. This essentially assists the educator with giving an exact depiction of the information [3]. Because of the acknowledgment of the significance of measurements, it has brought about the development of the way that insights is a significant and compulsory science to be dominated for most trains, for example, in the field of schooling as recently portrayed [5]. The insight that measurements are challenging to arise among understudies who do not have a definite (non-numerical) educational background [7]. Merak assumes that statistics is a science that is full of counting processes similar to mathematics.

As a result of this view, some students then feel anxious when facing statistics courses. The level of anxiety in each individual can be different. Anxiety is a behavior that arises because of a situation that the person who experiences it considers endangers their psychological state [8]. Anxiety about statistics is a feeling experienced by a person when dealing with all things related to statistics, for example in collecting data to conducting data analysis [9]. This feeling only lasted on the things that made them uncomfortable with stats. In the end, this feeling affects student learning outcomes in statistics courses and other subjects related to statistics such as research methodology [10].

Statistical anxiety can be grouped into three [11], namely i) Dispositional; ii) Situational; and iii) Environmental factors. The span of the instructing and growing experience completed, the presence of earlier information about numeracy can be one of the variables causing factual tension [12]. Initial knowledge of arithmetic with the intensity of learning statistics will be directly proportional to one's statistical anxiety. In addition, the age factor will also affect the emergence of one's anxiety [13]. Those who have a relatively old age will have excessive anxiety when compared to those who have a relatively young age.

Knowledge of the basics of research has a very important position for students, both first, second and third strata. Scientific works in the form of theses, theses and dissertations are one of the main requirements that must be taken or written by students. Meanwhile, the basic knowledge of statistics is very important to enable students to make scientific works in the form of theses, theses and dissertations. But statistics classes have always given students in the social sciences a lot of anxiety. Most students choose these social science classes to avoid taking math or other math-related classes [14]. However, students who take the social sciences must face statistics courses. Students may be negatively affected by this statistical worry. There is a risk that students' statistics performance may suffer, and that they will acquire feelings of inadequacy and low self-efficacy in tasks requiring statistical reasoning [15], [16]. It is connected to success not just in research programs and statistics classes. Measurement of statistical anxiety needs to be done to see how students view statistics.

Several instruments on statistical anxiety have been developed in various parts of the world, some of which are the statistical anxiety rating scale (STARS) [17], statistics anxiety inventory (STAI) [18], statistics anxiety scale (SAS) [19], statistical anxiety measure [20], and statistics anxiety scale [21]. To date, no statistical anxiety instrument has been developed in the Indonesian version. The psychometric evaluation of the instrument is usually carried out using a classical test theory approach. In the world of education, this shift occurs not only for the purpose of developing teaching materials but also in evaluating the instruments to be developed or recalibrated which used to use the classical test theory approach, but now use a modern test theory approach (item responder theory or item response theory) Rasch model [22]. Analysis using the Rasch model approach gives users the flexibility to examine many things in the context of measuring an instrument [23]. Separate estimations are carried out so that the results of the psychometric measurement/evaluation produced can be used for various characteristics of any respondent [24], [25].

In addition to using the Rasch model approach [26], this psychometric evaluation also uses a confirmatory factor analysis (CFA) approach. This is done to evaluate the construct of the SAS instrument that has been calibrated into the Indonesian version [27]. The use of CFA provides empirical reinforcement regarding the structure of the model making up an instrument [28]. From this explanation, the combination of the Rasch model with CFA will produce a standard instrument from the SAS instrument. Determining the nature of the concept (dimensions and indicators) as well as the measurement features of the SAS questionnaire using contemporary psychometric analysis is the purpose of this work (Rasch and CFA).

2. RESEARCH METHOD

2.1. Respondent

Respondents in this study consisted of 1,651 who came from various universities in Indonesia. The sampling technique used was cluster random sampling technique. The recipients came from three regions in the territory of Indonesia which included three regions (West, Central, and East) of Indonesia. Characteristics of respondents as in Table 1.

2.2. Rasch model

The probabilistic model, developed by George Rasch in 1960, was a game-changer in the field of psychometry. The Rasch model is a popular choice for creating reliable and valid assessment tools in many domains, including education [29], [30]. For this reason, it is very appropriate to use the Rasch model as a tool to carry out psychometric evaluations in this study [31]. The SAS is an instrument developed using a polytomy (Ordinal) scale. The form of the scale will be converted into an interval scale using the Rasch model with the aim of increasing the accuracy of the measurements made [30], [32]. One of part from Rasch model was rating scale model (RSM) is a model for polytomous data [33], [34] with the (1).

$$ln\left(\frac{P_{ni(x=k)}}{P_{ni(x=k-1)}}\right) = \theta_n - \delta_i - \tau_k \quad (1)$$

where, θ is a parameter of the respondent's ability, δ is the item parameter. The parameter τ is symbolized by the category parameter. Because RSM is part of the Rasch model, various analytical assumptions also apply [35].

Table 1. Characteristics of respondents

Variable	Amount	Percentage
Man	457	27.46%
Woman	1,194	72.54%
Western Indonesia	922	55.71%
Central Indonesia	605	36.76%
Wester Eastern Indonesia	124	7.53%

2.3. Data analysis
2.3.1. Confirmatory factor analysis
Factor analysis is an analytical method to find out whether there is latent variables (variable who cannot be observed directly) which is the reason why a set of variables are correlated with each other [36]. An instrument's construct validity as a psychological test or assessment may be examined using CFA. The degree to which all test items actually measure/provide information about just one thing, namely what is being measured, may be validated by CFA [37].

2.3.2. Rasch model
The Rasch model analysis was carried out using the RSM. The use of this model as a result of the SAS instrument using a rating scale. RSM is used to calibrate all items on the SAS instrument which is carried out with the help of Winsteps (4.0) [38].

3. RESULTS AND DISCUSSION
3.1. Results
3.1.1. Confirmatory factor analysis
Factor analysis is used to see the first order model. The results of the analysis show that the chi-square ($\chi 2$) is 468.79 with a p-value of less than 0.001. For the RMSEA value of 0.082 with 90% CI value moving from 0.637 to 0.930. Meanwhile, the CFI value is 0.975; SRMR value of 0.0472. All items have a significant factor loading (0.611 to 0.830) with a significance value of $p <0.05$.

3.1.2. Dimensionality
Unidimensional assumption is one of the requirements in the RSM analysis [39], [40]. The unidimensional criteria can be seen from the value of Raw variance explained by measures of more than 30% [41]. Based on the findings of the unidimensional analysis presented in Table 2, it can be observed that the results of the analysis show the value of Raw variance explained by measures 50.4%. This value is more than the predetermined criterion value so that the instrument test meets the unidimensional criteria. In addition, in another opinion, the eigenvalue of the Unexplained variance in 1st contrast has an observed percentage value of 6.2%.

3.1.3. Local independence
Local independence testing aims to ensure that the items and respondents are not interrelated, in other words, the difficulty level of an item does not depend on the characteristics of the respondent [42]. The test was carried out using the Q3 test [43]. From the results of the analysis used, it shows that the residual correlation of each item is not more than 0.30 which means that there is no interdependence of each item.

3.1.4. Items fit
In the Rasch model with the RSM approach, the position of item fit aims to find out which items match the model used, in this case the RSM model. This value indicates the function of the item to measure information from the measured variable which in this case is to measure statistical anxiety. To see whether the items fit or not,

the mean square (MNSQ) outfit value is in the range of 0.5 to 1.5 [44]. From the results of the analysis, it was found that the MNSQ outfit value for the SAS instrument was in the range of 0.64 to 1.47 as shown in Table 3.

Table 2. Unidimensionality analysis

Element	Eigen value	Observation		Expected value
Total raw variance in observations	34.2550	100.0%		100.0%
Raw variance explained by measures	17.2550	50.4%		50.2%
Raw variance explained by persons	10.6795	31.2%		31.1%
Raw Variance explained by items	6.5756	19.2%		19.1%
Raw unexplained variance (total)	17.0000	100.0%		100.0%
Unexplained variance in 1st contrast	2.1132	6.2%	12.4%	
Unexplained variance in 2nd contrast	1.9137	5.6%	11.3%	
Unexplained variance in 3rd contrast	1.5899	4.6%	9.4%	
Unexplained variance in 4th contrast	1.3901	4.1%	8.22%	
Unexplained variance in 5th contrast	1.2556	3.7%	7.4%	

Table 3. Item measure order

Items	Measure	Outfit MNSQ
14	0.5	0.80
6	0.48	0.73
17	0.47	0.64
16	0.45	0.73
15	0.39	0.70
8	0.36	1.34
4	0.22	1.30
7	0.15	1.41
2	0.01	0.87
5	-0.05	1.47
3	-0.05	0.90
10	-0.21	0.87
12	-0.29	0.94
11	-0.34	0.88
13	-0.64	0.72
9	-0.65	1.15
1	-0.79	1.08

Table 3 shows the logit and Outfit MNSQ values of the SAS instrument. It can be seen that the logit value moves from -0.79 to 0.5. Item 1 has a logit value of -0.79 which indicates that the item is very easy for respondents to answer. In contrast to item 14 which has a logit value of 0.5. The logit value is the highest from the other items, which indicates that the item is very difficult to answer by the respondent.

3.1.5. Rating scale diagnostic

In the form of an instrument with a response choice in the form of a rating scale, it certainly has an erratic interval from each answer choice. This means that the answers from each respondent, even though they are in the same category, may have different levels of similarity. In the measurement of the Rasch RSM, the form of analysis used to evaluate how well the SAS questionnaire functions to create a measure that can be interpreted well [45]. We conducted an analysis on each item, assessing the quantity of endorsements, the distribution pattern of endorsements, and the MNSQ statistics as shown in Table 4.

Table 4 shows the value of the diagnostic rating scale. From these results, it can be seen that the MNSQ infit value is not more than 2 [46]. The average observation value starts from -2.79 logit for the Strongly Disagree category; -1.22 logit for Disagree category; -0.15 for the Occasional category; 1.05 logit for the Agree category; and 2.11 logit for the Strongly Agree category. Figure 1 shows a graph of the answer choice categories by forming a recommended pattern where each category intersects each other.

Table 4. Diagnostic rating scale

Category	Threshold	Observed		Observed average	Infit MNSQ	Outfit MNSQ
		Count	%			
Strongly disagree	none	452	6	-2.79	1.18	1.19
Disagree	-3.31	1868	24	-1.22	0.91	0.90
Sometimes	-1.15	3115	40	-0.15	0.84	0.85
Agree	.81	2029	26	1.05	0.91	0.92
Strongly agree	3.65	288	4	2.11	1.58	1.48

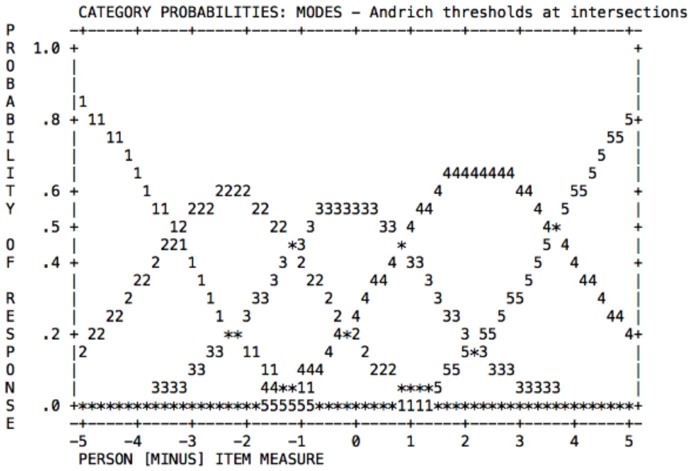

Figure 1. Category response curve

3.1.6. Reliability

The ability of an instrument's items to differentiate between test takers' skill levels is measured by a statistic called item separation. Reliability estimation is conducted for both individuals and things within the framework of Rasch RSM. Table 5 shows the person separation value of 3.40 with a reliability value of 0.92. A large separation value indicates that the instrument compiled is able to identify items (easy or difficult) and respondents (not able or able). Meanwhile, summary item (reliability and separation) is shown in Table 6.

Table 5. Summary Person

	Total Score	Count	Measure	Model S.E.	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	50.0	17.0	- 0.17	0.37	1.00	- 0.5	0.99	- 0.5
P.SD	10.9	0	1.52	0.06	0.76	2.3	0.76	2.3
S.SD	10.9	0	1.52	0.06	0.76	2.3	0.76	2.3
Max.	84.0	17.0	6.69	1.03	4.68	5.9	4.57	5.9
Min.	18.0	17.0	- 6.17	0.35	0.06	- 5.2	0.06	- 5.2
Real RMSE 0.43			True SD 1.46		Separation 3.40		Person Reliability 0.92	
Model RMSE 0.38			True SD 1.47		Separation 3.89		Person Reliability 0.94	
S.E. of Item Mean = 0.07								

Table 6. Summary item

	Total Score	Count	Measure	Model S.E.	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	1358.2	1651.0	0.00	0.07	1.00	- 0.5	0.99	- 0.5
P.SD	83.0	0	0.42	0.00	0.29	4.3	0.30	4.3
S.SD	85.5	0	0.43	0.00	0.30	4.5	0.31	4.4
Max.	1514.0	1651.0	0.50	0.07	1.61	8.0	1.66	8.3
Min.	1259.0	1651.0	- 0.79	0.07	0.64	- 6.4	0.64	- 6.4
Real RMSE 0.08			True SD 0.41		Separation 5.46		Item Reliability 0.97	
Model RMSE 0.07			True SD 0.41		Separation 5.80		Item Reliability 0.97	
S.E. of Item Mean = 0.10								

For the reliability and separation values from the item aspect, they are 0.97 and 5.46, respectively, which indicates that psychometrically the quality of the SAS instrument calibrated into Indonesian is very good according to the criteria for the separation index where the separation value is less than 2 indicating incompetence of the instrument distinguishes high and low abilities [47]. The reliability values for both respondents and items indicate that the SAS instrument that is developed consistently will provide consistent value for money measurement results in various respondents' conditions [48]. Meanwhile, it is clear from the results of this separation that the built SAS instrument can effectively categorize respondents into different subsets [49].

3.1.7. Wright map

On the Wright Map of the SAS instrument, it can be seen that the items in the questions are easier to respond to by all respondents [50], [51]. The findings presented in Figure 2 demonstrate the output results obtained from the Wright Map analysis. Item 14 which is at the very top and close to the dividing line, is the hardest. Item 1, which is closest to the dividing line, is the easiest. On the Wright Map, it can be seen that the difficulty level of the items has a not too large range (-0.79 to 0.5) while the respondent's ability has a fairly large range (-3.43 to 2.74).

```
MEASURE    PERSON - MAP - ITEM
               <more>|<rare>
   7               .  +
                      |
                   .  |
                      |
   6                  +
                   .  |
                      |
                      |
   5                  +
                      |
                   .  |
                      |
   4                  +
                      |
                   .  |
                   .  |
   3               .  +
                   . T|
                   #  |
                 .##  |
   2             .##  +
                 .##  |
               #####  |
               ##### S|
   1           .####  +
                .###  |T
            .#######  |S 14      15      16      17      6
             .######  |  4       7       8
   0      ##########  +M 2       3       5
      .############ M|  10      11      12
            .#######  |S
             .######  |T 1       13      9
  -1        .#######  +
              .#####  |
               .####  |
                 ### S|
  -2             .##  +
               .####  |
                  .#  |
                   .  |
  -3               .  +
                   . T|
                   #  |
                   .  |
  -4                  +
                      |
                   .  |
                      |
  -5               .  +
                      |
                   .  |
                      |
  -6                  +
                   .  |
                      |
                      |
  -7                  +
               <less>|<freq>
```

Figure 2. Wright map

3.1.8. Item information function (IIF)

Item Information Functions is a way to describe how hard an item on a test set is, choose test items, and compare two or more test sets. With the item information function, you can find out which item fits the model, which helps you choose which items to put on the test. The test information function is the sum of the information functions of the items that make up the test. As shown in Figure 3, the graph of the item information function tends to show that the information function is high.

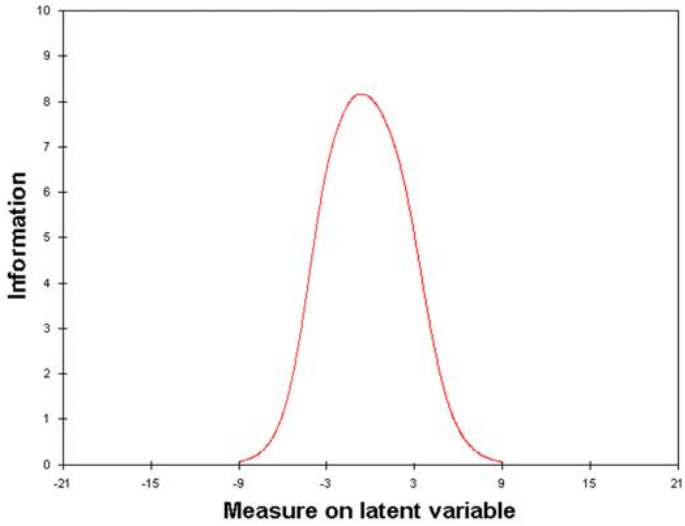

Figure 3. IIF SAS instrument

3.2. Discussion

This study uses a combination of two forms of data analysis that are classified as complex, namely the analysis of the Rasch and CFA models. Both analyzes were used to validate the developed SAS instrument. The results showed that all the items on the SAS instrument in general had fulfilled all the indicators of the validity of an instrument. The importance of testing the prerequisites for these unidimensional assumptions. One method is CFA. Using CFA to figure out the construct validity of the developed SAS instrument [52]. It's shows that the suitability of the SAS instrument's parts is based on the value of the Goodness of Fit indicator, which has already been said. Confirmatory factor analysis is useful for ensuring that all items on an instrument match the dimensions/aspects that form the basis of an instrument. Furthermore, all items on the SAS instrument have MNSQ outfit values according to the criteria (in the range of 0.5 to 1.5), which means that in general the items on the SAS instrument have consistency in measuring from one construct which underlie. This means that, on average, these questionnaire questions always measure the same thing [53]. This consistency can lead to precise measurements.

The findings of the item fit analysis may be used to learn more about the items that measure the generated construct, in this example, the statistical anxiety of a person. The ability to enhance the quality of these goods will be beneficial for people who take measurements in order to create reliable data [54]. In such conditions, it is necessary to do some treatment of items that are not appropriate, namely by conducting a review which will then be repaired or even replaced with new items. This is to ensure that in the future, measurements made regarding perceived statistical anxiety are the result of quality items.

On the Diagnostic Rating Scale of the SAS instrument, it shows that all categories of existing responses ranging from strongly disagree to strongly agree have consistently increased from low to high. This increase is in line with changes in the existing response categories. In addition, the MNSQ Outfit values from the five response categories are all included in the MNSQ Outfit value research category. This is important to ensure considering the position of the SAS instrument which aims to detect statistical anxiety will be very influential when the respondent cannot choose the correct response provided [55]. When respondents have difficulty in determining their answer choices, it is certain that the results of the statistical anxiety measurement will be biased.

In the Rasch model analysis using Winsteps, we can find out the distribution between the respondents' abilities and the quality of the items on the same diagram (using the same scale, namely the logit scale). The similarity of the scale used allows us to directly compare the two. The Wright Map enables researchers to determine how questionnaire questions generate latent variables by first examining an instrument's strengths and limitations, then recording the hierarchy of questionnaire items, and then comparing theory to observational data findings. From the results of the analysis obtained, it can be seen that all items can be answered well by the respondents as seen from the position of the items that are around the average ability of the respondents. Researchers hope that the results of this study will help them find people who have statistical anxiety and figure out how to help them.

4. CONCLUSION

From the results of the analysis, it shows that the SAS instrument developed based on the characteristics of learning in Indonesia shows good results as seen from the output of the analysis of the Rasch and CFA models. There is not a single item of statement that can give the respondent confusion in responding. From the results of the research conducted, it can be the first step for educational practitioners, especially for those who want to know about statistical anxiety so that they can prepare for the next step to overcome this anxiety, considering that statistics is one of the important sciences in education. The research can contribute to the development of the world of education, especially in diagnosing statistical anxiety in students in Indonesia. The results of this study will make it easier for people making decisions in the future to use the statistical anxiety questionnaire in educational evaluation studies. Even though the current study is about statistical anxiety questionnaires, the same method can be used to evaluate and validate other educational performance rating scales. The purpose of this research was to use contemporary psychometric analysis, such as the CFA, and independent item and sample models, such as the Rasch model, to determine the dimensions or component structure and measurement characteristics of the statistical anxiety questionnaire. We hope that this statistical anxiety questionnaire can then be used directly by lecturers to measure students' perceptions of their level of anxiety towards statistical anxiety without the need to pay attention to the validity, reliability, and interpretive quality of scores from the statistical anxiety questionnaire.

REFERENCES

[1] M. C. Michel, T. J. Murphy, and H. J. Motulsky, "New author guidelines for displaying data and reporting data analysis and statistical methods in experimental biology," *Molecular Pharmacology*, vol. 97, no. 1, pp. 49–60, Jan. 2020, doi: 10.1124/mol.119.118927.

[2] P. Mishra, C. Pandey, U. Singh, A. Gupta, C. Sahu, and A. Keshri, "Descriptive statistics and normality tests for statistical data," *Annals of Cardiac Anaesthesia*, vol. 22, no. 1, 2019, doi: 10.4103/aca.ACA_157_18.

[3] P. Steinberger, "Assessing the statistical anxiety rating scale as applied to prospective teachers in an Israeli teacher-training college," *Studies in Educational Evaluation*, vol. 64, Mar. 2020, doi: 10.1016/j.stueduc.2019.100829.

[4] D. Disman, M. Ali, and M. Syaom Barliana, "The use of quantitative research method and statistical data analysis in dissertation: an evaluation study," *International Journal of Education*, vol. 10, no. 1, pp. 46–52, Sep. 2017, doi: 10.17509/ije.v10i1.5566.

[5] C. Ramirez, C. Schau, and E. Emmioğlu, "The importance of attitudes in statistics education," *Statistics Education Research Journal*, vol. 11, no. 2, pp. 57–71, Nov. 2012, doi: 10.52041/serj.v11i2.329.

[6] D. S. Silva, G. H. S. M. de Moraes, I. K. Makiya, and F. I. G. Cesar, "Measurement of perceived service quality in higher education institutions," *Quality Assurance in Education*, vol. 25, no. 4, pp. 415–439, Sep. 2017, doi: 10.1108/QAE-10-2016-0058.

[7] S. Lewthwaite and M. Nind, "Teaching research methods in the social sciences: Expert perspectives on pedagogy and practice," *British Journal of Educational Studies*, vol. 64, no. 4, pp. 413–430, Oct. 2016, doi: 10.1080/00071005.2016.1197882.

[8] O. Remes, C. Brayne, R. van der Linde, and L. Lafortune, "A systematic review of reviews on the prevalence of anxiety disorders in adult populations," *Brain and Behavior*, vol. 6, no. 7, Jul. 2016, doi: 10.1002/brb3.497.

[9] R. J. Nesbit and V. J. Bourne, "Statistics anxiety rating scale (STARS) use in psychology students: a review and analysis with an undergraduate sample," *Psychology Teaching Review*, vol. 24, no. 2, pp. 101–110, 2018.

[10] T. Tutkun, "Statistics anxiety of graduate students," *International Journal of Progressive Education*, vol. 15, no. 5, pp. 32–41, 2019.

[11] H. S. A. Hamid and M. K. Sulaiman, "Statistics anxiety and achievement in a statistics course among psychology students," *International Journal of Behavioral Science*, vol. 9, no. 1, pp. 55–66, 2015.

[12] M. Paechter, D. Macher, K. Martskvishvili, S. Wimmer, and I. Papousek, "Mathematics anxiety and statistics anxiety. Shared but also unshared components and antagonistic contributions to performance in statistics," *Frontiers in Psychology*, vol. 8, Jul. 2017, doi: 10.3389/fpsyg.2017.01196.

[13] R. van der Cruijsen, J. Murphy, and G. Bird, "Alexithymic traits can explain the association between puberty and symptoms of depression and anxiety in adolescent females," *PLOS ONE*, vol. 14, no. 1, Jan. 2019, doi: 10.1371/journal.pone.0210519.

[14] Z. Wang *et al.*, "Is math anxiety always bad for math learning? The role of math motivation," *Psychological Science*, vol. 26, no. 12, pp. 1863–1876, Dec. 2015, doi: 10.1177/0956797615602471.

[15] S. Malik, "Undergraduates' statistics anxiety: A phenomenological study," *The Qualitative Report*, vol. 20, no. 2, pp. 120–133, Feb. 2015, doi: 10.46743/2160-3715/2015.2101.

[16] J. Jordan, G. McGladdery, and K. Dyer, "Dyslexia in higher education: Implications for maths anxiety, statistics anxiety and psychological well-being," *Dyslexia*, vol. 20, no. 3, pp. 225–240, Aug. 2014, doi: 10.1002/dys.1478.

[17] R. J. Cruise, R. W. Cash, and D. L. Bolton, "Development and validation of an instrument to measure statistical anxiety," in *American Statistical Association 1985 Proceedings of the Section on Statistical Education*, 1985, pp. 92–97.

[18] M. Zeidner, "Statistics and mathematics anxiety in social science students: some interesting parallels," *British Journal of Educational Psychology*, vol. 61, no. 3, pp. 319–328, Nov. 1991, doi: 10.1111/j.2044-8279.1991.tb00989.x.

[19] T. B. Pretorius and A. M. Norman, "Psychometric data on the statistics anxiety scale for a sample of South African students," *Educational and Psychological Measurement*, vol. 52, no. 4, pp. 933–937, Dec. 1992, doi: 10.1177/0013164492052004015.

[20] M. S. Earp, *Development and validation of the statistics anxiety measure*. Colarado: University of Denver, 2007.

[21] A. Vigil-Colet, U. Lorenzo-Seva, and L. Condon, "Development and validation of the statistical anxiety scale," *Psicothema*, vol. 20, no. 1, pp. 178–180, 2008.

[22] J. Kean, E. F. Bisson, D. S. Brodke, J. Biber, and P. H. Gross, "An introduction to item response theory and Rasch analysis: application using the eating assessment tool (EAT-10)," *Brain Impairment*, vol. 19, no. 1, pp. 91–102, Mar. 2018, doi: 10.1017/BrImp.2017.31.

[23] D. Andrich and I. Marais, *A course in Rasch measurement theory*. Singapore: Springer Nature Singapore, 2019. doi: 10.1007/978-981-13-7496-8.

[24] F. Fortuna and F. Maturo, "K-means clustering of item characteristic curves and item information curves via functional principal component analysis," *Quality & Quantity*, vol. 53, no. 5, pp. 2291–2304, Sep. 2019, doi: 10.1007/s11135-018-0724-7.

[25] S. E. Stemler and A. Naples, "Rasch measurement v. item response theory: knowing when to cross the line," *Practical Assessment, Research, and Evaluation*, vol. 26, 2021.

[26] C. Allison, S. Baron-Cohen, M. H. Stone, and S. J. Muncer, "Rasch modeling and confirmatory factor analysis of the systemizing quotient-revised (SQ-R) scale," *The Spanish Journal of Psychology*, vol. 18, Mar. 2015, doi: 10.1017/sjp.2015.19.

[27] D. B. Flora and J. K. Flake, "The purpose and practice of exploratory and confirmatory factor analysis in psychological research: Decisions for scale development and validation," *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, vol. 49, no. 2, pp. 78–88, Apr. 2017, doi: 10.1037/cbs0000069.

[28] T. A. Brown, *Confirmatory factor analysis for applied research* (2nd ed.). New York: Guilford Press, 2015.

[29] C. Van Zile-Tamsen, "Using Rasch analysis to Inform rating scale development," *Research in Higher Education*, vol. 58, no. 8, pp. 922–933, Dec. 2017, doi: 10.1007/s11162-017-9448-0.

[30] M. Planinic, W. J. Boone, A. Susac, and L. Ivanjek, "Rasch analysis in physics education research: Why measurement matters," *Physical Review Physics Education Research*, vol. 15, no. 2, Jul. 2019, doi: 10.1103/PhysRevPhysEducRes.15.020111.

[31] B. Sumintono, "Rasch model measurements as tools in assessment for learning," in *Proceedings of the 1st International Conference on Education Innovation (ICEI 2017)*, 2018, pp. 38–42. doi: 10.2991/icei-17.2018.11.

[32] W. J. Boone and J. R. Staver, *Advances in Rasch analyses in the human sciences*. Cham: Springer International Publishing, 2020. doi: 10.1007/978-3-030-43420-5.

[33] D. Andrich, "Rasch rating-scale model," in *Handbook of Item Response Theory*, Chapman and Hall/CRC, 2018, pp. 75–94.

[34] M. Tabatabaee-Yazdi, K. Motallebzadeh, H. Ashraf, and P. Baghaei, "Development and validation of a teacher success questionnaire using the Rasch model," *International Journal of Instruction*, vol. 11, no. 2, pp. 129–144, 2018.

[35] A. Dehqan, F. Yadegari, A. Asgari, R. C. Scherer, and P. Dabirmoghadam, "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*, vol. 31, no. 1, pp. 113.e19-113.e29, Jan. 2017, doi: 10.1016/j.jvoice.2016.03.018.

[36] H. W. Marsh, J. Guo, T. Dicke, P. D. Parker, and R. G. Craven, "Confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM), and set-ESEM: Optimal balance between goodness of fit and parsimony," *Multivariate Behavioral Research*, vol. 55, no. 1, pp. 102–119, Jan. 2020, doi: 10.1080/00273171.2019.1602503.

[37] A. G. Yong and S. Pearce, "A beginner's guide to factor analysis: Focusing on exploratory factor analysis," *Tutorials in Quantitative Methods for Psychology*, vol. 9, no. 2, pp. 79–94, Oct. 2013, doi: 10.20982/tqmp.09.2.p079.

[38] J. M. Linacre, *Winsteps Rasch measurement computer program user's guide*. Beaverton, Oregon: Winsteps.com, 2016.

[39] S. Lambert *et al.*, "Using confirmatory factor analysis and Rasch analysis to examine the dimensionality of The Patient Assessment of Care for Chronic Illness Care (PACIC)," *Quality of Life Research*, vol. 30, no. 5, pp. 1503–1512, May 2021, doi: 10.1007/s11136-020-02750-9.

[40] M. Geramipour, "Rasch testlet model and bifactor analysis: how do they assess the dimensionality of large-scale Iranian EFL reading comprehension tests?" *Language Testing in Asia*, vol. 11, no. 1, Dec. 2021, doi: 10.1186/s40468-021-00118-5.

[41] W. Rahayu, M. D. K. Putra, D. Iriyadi, Y. Rahmawati, and R. B. Koul, "A Rasch and factor analysis of an Indonesian version of the student perception of opportunity competence development (SPOCD) questionnaire," *Cogent Education*, vol. 7, no. 1, Jan. 2020, doi: 10.1080/2331186X.2020.1721633.

[42] R. Debelak and I. Koller, "Testing the local independence assumption of the Rasch model with Q3-based nonparametric model tests," *Applied Psychological Measurement*, vol. 44, no. 2, pp. 103–117, Mar. 2020, doi: 10.1177/0146621619835501.

[43] K. B. Christensen, G. Makransky, and M. Horton, "Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations," *Applied Psychological Measurement*, vol. 41, no. 3, pp. 178–194, May 2017, doi: 10.1177/0146621616677520.

[44] T. G. Bond and C. M. Fox, *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press, 2013.

[45] C. Jong, T. E. Hodges, K. D. Royal, and R. M. Welder, "Instruments to measure elementary preservice teachers' conceptions: an application of the Rasch rating scale model," *Educational Research Quarterly*, vol. 39, no. 1, pp. 21–48, 2015.

[46] J. M. Linacre, "Predicting responses from Rasch measures," *Journal of Applied Measurement*, vol. 11, no. 1, pp. 1–10, 2010.

[47] V. Aryadoust, L. Y. Ng, and H. Sayama, "A comprehensive review of Rasch measurement in language assessment: Recommendations and guidelines for research," *Language Testing*, vol. 38, no. 1, pp. 6–40, Jan. 2021, doi: 10.1177/0265532220927487.

[48] H. Abdullah, N. Arsad, F. H. Hashim, N. A. Aziz, N. Amin, and S. H. Ali, "Evaluation of students' achievement in the final exam questions for microelectronic (KKKL3054) using the Rasch model," *Procedia - Social and Behavioral Sciences*, vol. 60, pp. 119–123, Oct. 2012, doi: 10.1016/j.sbspro.2012.09.356.

[49] H. Othman, I. Asshaari, H. Bahaludin, Z. M. Nopiah, and N. A. Ismail, "Application of Rasch measurement model in reliability and quality evaluation of examination paper for engineering mathematics courses," *Procedia - Social and Behavioral Sciences*, vol. 60, pp. 163–171, Oct. 2012, doi: 10.1016/j.sbspro.2012.09.363.

[50] F. N. Hikmah, M. I. Sukarelawan, T. Nurjannah, and J. Djumati, "Elaboration of high school student's metacognition awareness on heat and temperature material: Wright map in Rasch model," *Indonesian Journal of Science and Mathematics Education*, vol. 4, no. 2, pp. 172–182, Jul. 2021, doi: 10.24042/ijsme.v4i2.9488.

[51] D. C. Briggs, "Interpreting and visualizing the unit of measurement in the Rasch model," *Measurement*, vol. 146, pp. 961–971, Nov. 2019, doi: 10.1016/j.measurement.2019.07.035.

[52] L. A. Clark and D. Watson, "Constructing validity: New developments in creating objective measuring instruments," *Psychological Assessment*, vol. 31, no. 12, pp. 1412–1427, Dec. 2019, doi: 10.1037/pas0000626.

[53] C. M. Grilo, D. L. Reas, C. J. Hopwood, and R. D. Crosby, "Factor structure and construct validity of the eating disorder examination-questionnaire in college students: Further support for a modified brief version," *International Journal of Eating Disorders*, vol. 48, no. 3, pp. 284–289, Apr. 2015, doi: 10.1002/eat.22358.

[54] W. J. Boone, J. R. Staver, and M. S. Yale, *Rasch analysis in the human sciences*. Dordrecht: Springer Netherlands, 2014. doi: 10.1007/978-94-007-6857-4.

[55] H. S. You, K. Kim, K. Black, and K. W. Min, "Assessing science motivation for college students: Validation of the science motivation questionnaire II using the Rasch-Andrich rating scale model," *EURASIA Journal of Mathematics, Science and Technology Education*, vol. 14, no. 4, Jan. 2018, doi: 10.29333/ejmste/81821.

BIOGRAPHIES OF AUTHORS

Deni Iriyadi is a lecturer at the Postgraduate Program at the State Islamic University of Sultan Maulana Hasanuddin Banten, Serang, Banten, Indonesia. He was appointed as a lecturer at the university in 2020. Dr Deni's research interests lie in the areas of assessment in education, mathematics education, educational measurement, school-based assessment, and learning evaluation. He can be contacted via: deni.iriyadi@uinbanten.ac.id.

Ahsanul Khair Asdar is a lecturer at the Postgraduate Program at the Sriwijaya State Buddhist College Tangerang, Banten, Indonesia. He was appointed as a lecturer since 2018. His research interests lie in the areas of assessment in education, mathematics education, educational measurement, and learning evaluation. He can be contacted via email: ahsanul.khair@stabn-sriwijaya.ac.id.

Bambang Afriadi is a lecturer at Syekh Yusuf Islamic University Tangerang and is interested in research and evaluation methods. His publication interests lie in social education, education management, education evaluation, and school-based evaluation and learning. Several books and research results in scientific journals have been published in the education and social fields. He can be contacted via email: bambang.afriadi@unis.ac.id

Mohammad Ardani Samad is a lecturer at the Undergraduate Programme of Pelamonia Institute of Health Sciences Makassar, Makassar, South Sulawesi, Indonesia. He was appointed as a lecturer at the institute in 2015. Ardani's research interests lie in the field of mathematics education. He can be contacted at: ardani.samad@stikespelamonia.ac.id.

Yogi Damai Syaputra is a Guidance and Counseling Lecturer at Sultan Maulana Hasanuddin State Islamic University Banten, Indonesia. Research fields include educational assessment, guidance and counseling assessment, counseling techniques and skills, psychological assessment, social psychology. He can be contacted via email: yogi.damai@uinbanten.ac.id.