Does New Normal Learning Anxiety Scale (NNLAS) Worth to Measure Anxiety? A Study to Investigate Instrument Characteristic

Anggi Ristiyana Puspita Sari[1*], *Ruli* Meiliawati[1], *Deklin* Frantius[1], *Juliana* Sihombing[1], and *Gryshelda* Nabila[1]

[1]Chemiatry Education, Universitas Palangka Raya, 74874 Kota Palangka Raya, Indonesia

Abstract. During the COVID-19 pandemic, students developed study habits to adapt to online learning, enjoying discussing with lecturers and friends through meeting platforms. However, the lack of interaction gradually erodes their confidence. As a result, they become concerned about the new normal learning environment, where they will have to face lecturers and friends. They must be prepared and equipped to cope with this new environment. Building students' readiness and confidence in the new normal learning setting can reduce anxiety. Nevertheless, students experience anxiety when it comes to returning to face-to-face learning. Consequently, an attempt has been made to develop the New Normal Learning Anxiety Scale (NNLAS) questionnaire to assess and verify its characteristics. The model development of NNLAS is based on the 4D model. In order to establish content validity, the initial product was evaluated by five experts and explored using Aiken's formula. This research involved 209 undergraduates from Universitas Palangka Raya to assess construct validity. The Rasch model was employed to examine various characteristics of NNLAS, including unidimensionality, reliability, item difficulty, item fit, and rating scale. The results demonstrate that NNLAS is suitable for measuring students' learning anxiety. ***Keywords***: New Normal Learning Anxiety Scale (NNLAS), anxiety, Rasch model.

1 Introduction

The Covid-19 pandemic has established a new normal that affects everyone. As a result, the education sector is confronted with numerous challenges in delivering instruction to students. Education is learning skills to develop students' potential actively. The ideal learning process should maximize the role of lecturers as facilitators and evaluators while students as subjects who are involved in the learning process. If lecturers and students do not actively participate in the learning process, the learning outcomes will be less than optimal. However, during the Covid-19 pandemic, the world of education was disrupted due to the impact of the pandemic. Of course, the Covid-19 pandemic has limited the space for the social aspects of society[1].

The COVID-19 pandemic forces humans to live a life with limited mobility. In this condition, almost all activities can be done *online*. Undeniably, these conditions bring serious challenges to all sectors, including the world of education[2]. The Indonesian government issued various regulations to respond to the situation in the pandemic era. The Circular Letter of the Minister of Education and Culture Number 4 of 2020 concerning implementing Education in the Coronavirus Disease (COVID-19) Emergency Period states that all learning activities are done through online or distance learning. Therefore, education practitioners must work hard to arrange appropriate learning strategies.

The need to find and process information in learning in the pandemic era is a basic need that students must master. As a result of these conditions, students become more independent and skilled in using digital devices during the learning process so that students are able to carry out their roles as active learners. However, this condition also has an impact on the psychological condition of students. Students are used to online learning where lecturer supervision is not optimal. Meanwhile, online learning is not always carried out considering the conditions in this pandemic era are getting better [3] .

Some universities have implemented offline lectures although not yet fully. This forces students to get out of their comfort zone where they can no longer relax attending lectures from home or other places. Therefore, the offline lecture process triggers student learning worries and anxiety [4]. Student learning anxiety does not only occur during the lecture process in class, but also when carrying out practicum in the laboratory. Learning anxiety is a variable that has a negative effect on lectures [5,6]. If students have anxiety about chemistry lectures, it will make them lose interest in science.

Affective variables, especially learning anxiety, should not be ignored because they can stimulate cognitive knowledge and have an impact on learning outcomes. High levels of anxiety cause students to not be able to show their true level of proficiency and can interfere with performance. This chemical anxiety represents student feelings such as fear and anxiety about chemistry. Students with a negative attitude towardtudying chemistry are known as *chemophobia* or chemophobia[7]. In chemistry education, several studies have revealed that the problem of perception and attitude towards chemistry[8], student achievement and

* Corresponding author: anggi.ristiyana@fkip.upr.ac.id

performance in chemistry, and lecturer pedagogic skills correlateith learning anxiety[8].

Knowing the level and causes of anxiety will be effective in identifying ways to relieve anxiety and reorient students. Therefore, it is important to identify students' concerns. If this is ignored, students will not be able to show true proficiency and have an impact on learning outcomes that are not optimal. In addition, it is important to know the level of student chemistry learning anxiety and its causes so that lecturers can apply appropriate learning strategies when offline lectures will be carried out. Based on this explanation, this study aims to: 1) develop a *chemistry anxiety* and 2) measure the level of student chemistry learning anxiety. The items in the *chemistry anxiety* validated using factor analysis (EFA) and their characteristics were analyzed using Rasch modeling.

2 Literature Review

According to several research [9–11], Anxiety levels significantly impact learning. Anxiety acts as a detrimental factor that negatively affects the learning process. When engaging with scientific concepts, individuals who experience science anxiety often face intense fear and tension. Science anxiety refers to a specific fear of learning science. Chemophobia is the term used in chemical education to describe learning anxiety caused by chemistry lectures. Chemophobia can be defined as phobic anxiety associated with a fear of chemicals and a science-based fear of chemicals[12].

High anxiety students were considerably more likely to perform poorly academically than other students. Chemical anxiety mostly impacts a person's fear of chemistry lectures, particularly when studying chemistry, evaluative chemistry-related events, such as assessments or tests, and also experiencing fear when handling chemicals in a laboratory environment. Chemophobia is one of the learning obstacles that prevents pupils from developing into exceptional students in the chemical field[13].

The events of stress, worry, extreme fear, and anxiety when studying chemistry are also known as learning because a student who has a negative perspective of chemistry would feel bored, worried [14], fearful, and inferior during the learning process. Anxiety is a fundamental feeling that everyone experiences occasionally and is concerned about an uncertain future. On the other hand, crisiss an emotion that no one wants since it will obstruct their ability to think scientifically[15].

According to Kamarudin[16], anxiety about studying chemistry is brought on by a number of factors, including: (1) anxiety when completing math problems, tests, or evaluations; (2) teachers who do not comprehend the subject; (3) working in a chemical laboratory; (4) students' negative perceptions; and (5) when hearing the word "chemistry". According to several research (Eddy, 2020; Senocak & Baloglu, 2014), learning chemistry anxiety is brought on by three things: analyzing chemistry, handling chemicals, and studying chemistry. The extensive syllabus, insufficient understanding of employment options in chemistry, a lack of field trips, inadequate laboratory equipment, and poor teaching strategies are the main reasons of worry when studying chemistry. According to research findings, learning anxiety can negatively impact student learning outcomes in chemistry lessons, learning motivation, and ability to think clearly. It can also negatively impact academic achievement and academic achievement, learning outcomes in receiving chemistry lessons, and student learning outcomes.

Chemistry anxiety was used to quantify student learning anxiety and define its level. The measurement's goal is to get precise and pertinent data. According to[17], measuring is the quantitative application of established criteria to assign qualities to things in the form of numbers or scores. The classical test theory approach (Classical Test Theory, CTT) and item response theory are the two main types of estimation or approaches to empirical assessment in education (Item Response Theory, IRT). The limits of traditional test theory[18] the creation of the IRT approach. The IRT model is excellent for instrument development and can give error estimates for each student and item. An interpretation can be more precisely made by estimating each inaccuracy. The IRT 1-PL model and the Rasch model share similarities, according to[19] Rasch modeling, which is based on a systematic response pattern, has an advantage over other techniques, especially traditional test theory, in that it can anticipate missing data. Rasch modeling offers advantages that make it obvious that statistical analysis findings on student exam results are more precise.

3 Method

3.1 Type of Research

This research is a type of quantitative descriptive research that aims to analyse a symptom[20]. This study was conducted to determine the value of the independent variable without making comparisons or connecting with other variables.

3.2 Procedure

The development of this product follows the steps involved in creating social psychology instruments [21,22]. However, only the development and implementation stages have been conducted in this study. The evaluation and analysis of structural modeling will be carried out in the subsequent year's research.

1. Determining the construct and specification of the instrument
   The stage of determining the construct and specifications of this instrument comprises three parts: analysis, establishing conceptual and operational definitions, and identifying indicators.
   a. Analysis
      The analysis phase is a process of analysing targets, analysing needs (*needs assessment*), and analysing a problem to determine research objectives. The steps in this analysis phase include:
      i. The analysis of the instrument domain aims to prepare the non-test instrument domain

to be suitable for measuring *chemistry anxiety* in students. Thus, this non-test instrument can measure the expected ability.

ii. Target analysis, which is to determine the target that will use the product developed.

iii. Objective analysis is to formulate the objectives to be achieved with the research and determine the importance of developing the resulting product.

b. Determining the conceptual definition and operational definition
The conceptual definition is determined by looking for relevant theories from reference books, journal articles, and previous research. After determining the conceptual definition, it is further translated into an operational definition.

c. Develop instrument item indicators
Based on the conceptual and operational definitions that have been determined, it can be translated into several indicators.

2. Writing Instrument Items
Writing instrument items was constructed from predetermined indicators for *chemistry anxiety*. The instrument's items are then arranged using a Likert scale form with 4 answer choices, with a "1" code for strongly disagree to "4" code for strongly agree.
3. Instrument Item Review
After the completion of the statement items, the review of instrument items was conducted through *expert judgment*. This research involved expert lecturers who provided validation feedback, based on which revisions were made.
4. Conducting Trials
The trials were carried out after the instrument items were revised based on input from the experts. The instrument testing phase was conducted on students from the University of Palangka Raya.
5. Analysing Instrument Items
This stage was carried out to analyse the quality of the test items of validity and reliability.
6. Revising Instrument Items
Revision of instrument items is carried out based on the results of instrument item analysis.
7. Assembling the Final Items
At this stage, the instrument is reassembled. Items that are declared unfit are then revised/discarded.

3.3 Subject

Determining the number of samples needed to represent the population in a descriptive quantitative study using a questionnaire is crucial. The sampling method used in this study is the Krejcie and Morgan method [21]. The equation made by Krejcie and Morgan in determining the sample depends on the number of populations, the level of accuracy in the form of proportion, and the proportion of the population. The population in this study were all students of the University of Palangka Raya who studied chemistry. This study involved 209 students from various majors at the University of Palangka Raya in the pilot phase. Meanwhile, the measurement phase involved 108 Chemistry Education study program students. Meanwhile, the object of this research is students' anxiety level in studying chemistry in offline lectures.

3.4 Research Instrument

The research instrument used was *chemistry anxiety* developed based on a predetermined development procedure. The instrument was then used to explore its characteristics.

3.5 Data Analysis Technique

3.5.1 Content Validity

The content validity analysis in this study uses the Aiken formula to calculate the content-validity coefficient based on the assessment of n experts on an item in terms of the extent to which the item can represent the measured construct.

3.5.2 Construct Validity

Construct validity in this study was obtained from the data from field trials. Data were analysed using exploratory factor analysis (EFA) techniques using the SPSS program.

3.5.3 Reliability

Reliability analysis in this study was carried out using the Winsteps program. This is done because through Rasch modeling can be known person reliability and item reliability. Pearson reliability states the quality of the respondents as a whole, while item reliability states the quality of the items as a whole. Meanwhile, the Cronbach Alpha obtained shows the interaction between the person and the item as a whole.

3.5.4 Item Characteristic

Analysis of item characteristic was carried out using Winsteps program. The item characteristic includes item difficulty, item fit, and rating scale.

4 Result and Discussion

4.1 Product Development

The product developed in this study was the *New Normal Learning Anxiety Scale* (NNLAS) questionnaire used to measure students' learning anxiety levels. According to experts, this questionnaire was compiled based on a grid developed from various aspects. This questionnaire consists of 26 statement items described from 2 aspects and 4 indicators. The results of the development progress through the following stages:

4.1.1 Results of Questionnaire Domain Analysis, Purpose and Objectives

The analysis of the questionnaire domain aimed to structure it in a way suitable for measuring students' levels of learning anxiety. Simultaneously, target analysis was conducted to examine the characteristics of students. The target group consisted of students who had experienced online learning during the lecture process

and transitioned to offline learning. Goal analysis was conducted to formulate the purpose of developing a questionnaire specifically designed to measure students' learning anxiety.

4.1.2 Product Target Results

At this stage, the questionnaire construct was designed. The design of the NNLAS questionnaire was prepared based on the analysis of operational definitions from various experts. At this stage, a questionnaire grid is designed consisting of aspects, operational definitions, and indicators. In addition, the determination of the number of items for each indicator is also carried out at this stage. Based on the two aspects developed by the experts, two indicators were obtained each and further elaborated into several sub-indicators. The initial product developed contained a total of 26 statements consisting of 14 favorable statements and 12 unfavorable statements.

4.2 Product Trial

Questionnaire trials were conducted during the second to fourth week of August. Questionnaires were distributed to various study programs and majors at the University of Palangka Raya, involving a total of 209 students. Product trials were conducted with the aim of obtaining instrument characteristics in the form of content validity from *expert judgment*, construct validity, reliability, item fit (*item fit*) with the model, item information function, item *misfit* (*item fit order*), respondent logit value (*person measure*), and analysis of the validity of the rating scale.

4.2.1 Content Validity

At the validation stage by *expert judgment*, quantitative data was obtained based on a validation questionnaire sheet. Analysis of the validity of the instrument used for this study uses the Aiken formula to calculate the content-validity coefficient based on the assessment results of 8 expert judgments that have been appointed. Miller [22] revealed thait was analyzed using the Aiken's V index to prove content validity through a Likert scale. The results of the Aiken's V index for each item are presented in Table 1.

Table 1. Aiken's V value for each item

Item	Aiken's V Index	Item	Aiken's V Index
Item 1	1,00	Item 14	0,93
Item 2	0,87	Item 15	0,87
Item 3	0,93	Item 16	0,93
Item 4	0,93	Item 17	0,87
Item 5	0,87	Item 18	0,93
Item 6	0,87	Item 19	0,93
Item 7	0,87	Item 20	0,93
Item 8	0,87	Item 21	0,93
Item 9	0,87	Item 22	0,87
Item 10	0,87	Item 23	0,93
Item 11	0,93	Item 24	0,87
Item 12	1,00	Item 25	1,00
Item 13	1,00	Item 26	1,00

4.2.2 Test Results of Rasch Modeling Assumptions

The validity of this construct shows the extent to which the instrument reveals a certain ability or theoretical construct to be measured. This construct validity analysis was obtained from the data from the field trials. Construct validity analysis uses *exploratory factor analysis* (EFA) and Rasch modeling techniques. The interpretation of this analysis will be used as a unidimensional assumption to determine the estimated number of dimensions that can be measured from the instruments that have been prepared. Unidimensional means that each item of the statement only measures one dimension. In this study, the assumed dimension is student learning anxiety.

Factor analysis begins by testing the adequacy of the sample used in the analysis, then the computer compiles a variance-covariance matrix, then calculates the eigenvalues. This eigenvalue is then used to calculate the percentage of explained variance an describe *the scree plot* [23]. The *factor analysis output* includes KMO statistics, Bartlett test, and *scree plot.* Table 2 shows the results of the KMO and Bartlett tests.

Table 2. Test Results of KMO and Bartlett

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0,769
Bartlett's Test of Sphericity	Approx. Chi-Square	2295,209
	df	325
	Sig	0,000

Table 2 shows that the sample size of 209 in this study is sufficient and is included in the good category because the KMO value obtained is greater than 0.5 and the Bartlett test has a significance of less than 0.05 (Beavers et al., 2013). Because these results are included in the good category, then to find out the unidimensional assumption, it can be analysed further using factor analysis.

The raw variance data output is 24.954%, indicating that the minimum requirement of 20% for unidimensionality has been met. As the unidimensional requirements have been satisfied, it can be concluded that the construct validation for the developed NNLAS

questionnaire has also been achieved. Additionally, Figure 1 illustrates the scree plot.

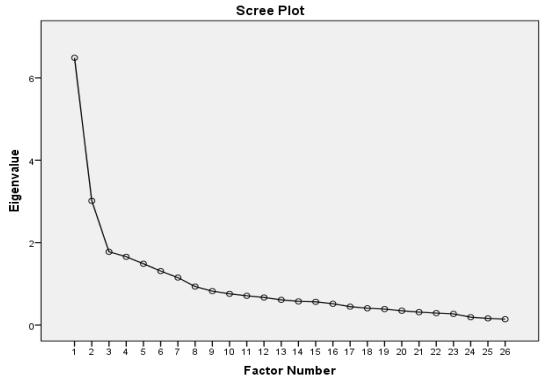

Fig. 1. Scree Plot Results of Exploratory Factor Analysis

In this study, the unidimensional assumption test was also analyzed using the Winsteps 3.73 program. The output of the unidimensional test is presented in Figure 2. Based on Figure 2, the raw data variance *is* also more than 20%, which is 39.5%. Unidimensionality of the instrument is an important measure to evaluate whether the developed instrument can measure what it is supposed to measure, in this study the dimension that was measured was learning anxiety. The analysis of the Rasch model uses principal component analysis of the residuals, which measures the extent to which the diversity of the instruments measures what should be measured [24].

```
TABLE 23.0 Item Dimensionality                ZOU748WS.TXT  Sep 4 14:38 2022
INPUT: 209 Person  26 Item  REPORTED: 209 Person  26 Item  4 CATS  WINSTEPS 3.73
-------------------------------------------------------------------------------

    Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
                                        -- Empirical --    Modeled
Total raw variance in observations   =        43.0 100.0%        100.0%
  Raw variance explained by measures =        17.0  39.5%         39.6%
    Raw variance explained by persons =         5.4  12.6%         12.6%
    Raw Variance explained by items   =        11.6  26.9%         27.0%
  Raw unexplained variance (total)    =        26.0  60.5% 100.0%   60.4%
    Unexplned variance in 1st contrast =         3.9   9.1%  15.0%
    Unexplned variance in 2nd contrast =         2.6   6.0%   9.9%
    Unexplned variance in 3rd contrast =         2.0   4.7%   7.8%
    Unexplned variance in 4th contrast =         2.0   4.6%   7.5%
    Unexplned variance in 5th contrast =         1.6   3.8%   6.3%
```

Fig. 2. Unidimensional assumption test results

4.2.3 Reliability

Summary statistics provide overall information about the quality of respondents as a whole, the quality of the instruments used, and the interaction between persons and items. Based on the result, the reliability value for the item is 0.98 while the reliability coefficient for the person 0.82. Meanwhile, Cronbach's Alpha 0.86. The reliability of Cronbach 's Alpha shows the interaction between person and item as a whole. If Cronbach's Alpha > 0.8,t can be categorized as very good [24]. From the value of person reliability of 0.82 and item reliability of 0.98, it can be concluded that the consistency of the answers of the respondents is good or strong enough and the quality of the items in the instrument is categorized as good.

4.2.4 Item Difficulty

Parameter item difficulty (δ) describes an item's approximate difficulty level expressed in log units. The index of difficulty level usually ranges from -2.0 logit to +2.0 logit (Hambleton & Swaminathan, 1985). Item difficulty level is one factor that influences the probability of a respondent's answer in response to a particular item. In the instrument in the form of a questionnaire, what is meant by difficult items are not items that are difficult to answer, but items that are the most difficult for respondents to agree with. The item difficulty level analysis results are presented in the form of output item measures.

Based on output, it can be observed that the range measures obtained was from -1.22 logit to +1.43 logit. In the results of their research, Adedoyin & Mokobi (2013) revealed that if the δ is between -0.5 to 0.5 logit, then the item is said to have a moderate or sufficient level of difficulty index. Meanwhile, if an item with a value of δ is below -1 logit indicates that the item is very easy, whereas if it is above +1 logit the item is said to be very difficult. Based on this explanation, item 7 with +1.43 logit shows the most difficult item for respondents to agree with, while item 13 with -1.22 logit is the easiest item for respondents to agree with. Thus, overall, the 26 item statements on the NNLAS instrument have a level of difficulty with an easy level.

4.2.5 Item Fit

Item fit is an index that determines whether an item functions optimally and meets the requirements as a good measurement tool or not. Sumintono & Widhiarso (2014) described that in order to be able to check for fit and misfit the INFIT MNSQ value of each item could be used. The calculation is done, namely the average value and standard deviation are added up and then compared. A logit value that is greater than this value indicates a misfit. Another criterion that can be used is the value outfit mean square (MNSQ) which is 0.5 < MNSQ < 1.5. The analysis results show that the items that are misfit are items 5 and 7 because they have a misfit greater than 1.5. Therefore, points 5 and 7 are declared invalid.

4.2.6 Rating Scale

The validity of the rating scale is a test conducted to verify whether the rating options used are confusing to respondents or not. Rasch model analysis provides a verification process for the rating assumptions given in the instrument. In Figure 3 it can be seen that the average observation starts from a logit of -0.68 for choice score 1 (strongly disagree) and finally increases to +1.37 logit for choice with a score of 4 (strongly agree).

```
TABLE 3.2 Rating Scale                        ZOU748WS.TXT  Sep 5 14:38 2022
INPUT: 209 Person  26 Item  REPORTED: 209 Person  26 Item  4 CATS  WINSTEPS 3.73
-------------------------------------------------------------------------------


SUMMARY OF CATEGORY STRUCTURE.  Model="R"
-------------------------------------------------------------------------------
|CATEGORY   OBSERVED|OBSVD SAMPLE|INFIT OUTFIT|| ANDRICH |CATEGORY|
|LABEL SCORE COUNT %|AVRGE EXPECT|  MNSQ  MNSQ||THRESHOLD| MEASURE|
|-------------------+------------+------------++---------+--------|
|  1   1     581  11|  -.68  -.72|  1.06  1.16||  NONE   |( -2.36)| 1
|  2   2    1002  18|  -.12  -.05|   .89   .85||   -.94   |   -.83 | 2
|  3   3    2339  43|   .70   .68|   .87   .86||   -.53   |    .64 | 3
|  4   4    1512  28|  1.37  1.38|  1.09  1.07||   1.47   |(  2.66)| 4
-------------------------------------------------------------------------------
OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.
```

Fig. 3. Results Output Rating Scale

Based on the Figure 2, there is no increase in the logit value. This shows that respondents could not make sure of choices 1 (strongly disagree) and 2 (disagree). Another measure that can be used is the Andrich Treshold to test whether the polytomous value used is correct or not. The Andrich Treshold moves from NONE, then negative, then continues to positive, sequentially indicating that the options given are valid for the respondent. Examining category thresholds involves the inspection of category probability curves to determine if the response probabilities are arranged in ascending order concordant with the categories, which would indicate ordered thresholds. The reliability and validity indicators of resulting category thresholds should then be assessed in order to evaluate how the new rating scale is functioning overall[25].

5 Conclusion

NNLAS was successfully developed through the steps of developing social psychology instruments. The Rasch model was conducted to explore its characteristic include unidimensionality, reliability, item difficulty, item fit, and rating scale. The results show that NNLAS is suitable for measuring students' learning anxiety.

References

1. E. Kiran, International Journal on New Trends in Education and Their Implications **3**, 90 (2012)
2. O. B. Adedoyin and E. Soykan, Interactive Learning Environments **28**, 1 (2020)
3. A. Haleem, M. Javaid, M. A. Qadri, and R. Suman, Sustainable Operations and Computers **3**, (2022)
4. Regita Septyani Rahmi and B. Murtafi'ah, Eduvelop: Journal of English Education and Development **6**, (2022)
5. C. Aydoğdu, International Journal of Progressive Education **13**, 85 (2017)
6. W. Pangestika and A. Wiyarsi, in *Advances in Social Science, Education and Humanities Research* (2021), pp. 241–247
7. R. M. Eddy, J Chem Educ **77**, (2000)
8. W. Pangestika and A. Wiyarsi, in *Advances in Social Science, Education and Humanities Research* (2021), pp. 241–247
9. E. Kaya and P. S. Çetin, International Journal on New Trends in Education and Their Implications **3**, 90 (2012)
10. J. Mallow, H. Kastrup, F. B. Bryant, N. Hislop, R. Shefner, and M. Udo, J Sci Educ Technol **19**, 356 (2010)
11. C. Aydoğdu, International Journal of Progressive Education **13**, 85 (2017)
12. R. M. Eddy, Journal of Chemistry Education **77**, 514 (2020)
13. M. Ajmal and S. Ahmad, Bulletin of Education and Research **41**, 67 (2019)
14. E. Kaya and P. S. Çetin, International Journal on New Trends in Education and Their Implications **3**, 90 (2012)
15. N. F. Kamarudin, N. H. Ibrahim, J. Surif, M. Ali, C. A. Talibm, and N. I. Ismail, International Recent Technology and Engineering **7**, 725 (2019)
16. N. F. Kamarudin, N. H. Ibrahim, J. Surif, M. Ali, C. A. Talibm, and N. I. Ismail, International Recent Technology and Engineering **7**, 725 (2019)
17. Miller, *Test Validation: A Literature Review* (University of Florida, Florida, 2003)
18. R. K. Hambleton and H. Swaminathan, *Item Response Theory: Principles and Applications*, 1st editio (Springer Science+Bussiness Media, New York, 1985)
19. B. Sumintono and W. Widhiarso, *Aplikasi Pemodelan Rasch Pada Asesmen Pendidikan* (Trim Komunikata, Cimahi, 2015)
20. J. W. Creswell, Research Design Qualitative Quantitative and Mixed Methods Approaches (2014)
21. B. Sumintono and W. Widhiarso, *Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial* (Trim Komunikata, Cimahi, 2014)
22. Miller, *Test Validation: A Literature Review* (University of Florida, Florida, 2003)
23. H. Retnawati, *Validitas, Reliabilitas, Dan Karakteristik Butir* (Parama Publishing, Yogyakarta, 2016)
24. B. Sumintono and W. Widhiarso, *Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial* (Trim Komunikata, Cimahi, 2014)
25. M. Robinson, A. M. Johnson, D. M. Walton, and J. C. Macdermid, BMC Med Res Methodol **19**, 1 (2019)