JPPIPA 8(5) (2022)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Analysis of the Characteristics of the Ethnoscience-Based Numeracy Test Instrument Using the Rasch Model

Muliani[1], Nanda Novita[1*], Mellyzar[2], Ali Imron Pasaribu[2], Mhd. Ridwan Fadli[2]

[1] Department of Physics Education, Universitas Malikussaleh, North Aceh, Indonesia
[2] Department of Chemistry Education, Universitas Malikussaleh, North Aceh, Indonesia.

Received: August 28, 2022
Revised: November 13, 2022
Accepted: November 20, 2022
Published: November 30, 2022

Corresponding Author:
Nanda Novita
nanda.novita@unimal.ac.id

DOI: 10.29303/jppipa.v8i5.2285

Abstract: The numeracy ability of students in Indonesia is in a low category based on PISA results from 2000 to 2018. So, efforts are needed to strengthen numeracy skills through the development of ethnoscience-based numeracy test instruments. The purpose of this study was to analyze the characteristics of the ethnoscience-based numeracy test instrument using the *Rasch model*. The subjects of this study were students of SMPN Arun and SMPS Sukma Bangsa, Lhokseumawe City, totaling 123 students. The number of questions to be analyzed in the test is 68 items. Analysis using *Winsteps software*. The results of the analysis showed that the ethnoscience-based numeration instrument in this study was declared reliable. 59 of the 68 items were said to be valid, with details of 32 valid items eligible to be used and 27 valid items needing to be revised for improvement before being used. This instrument can be used to measure subjects who have ethnoscience-based numeracy skills in the low, medium, and high categories, so to examine ethnoscience-based numeracy skills students can use this instrument.

Keywords: Ethnoscience; Numeracy; Test instrument; Rasch Model.

Introduction

The era in the 21st century makes development faster and more complex. In the era of the 21st century where science and technology, especially technology communication are growing very rapidly which has an impact on free competition that is so tight in all aspects of human life. In the tight challenges faced by society, a paradigm shift in the education system is needed that can provide a set of skills needed by students to face every aspect of global life (Karatas & Arpaci, 2021; Pratiwi et al., 2019; Widyastika et al., 2022).

The skills that students must have followed market needs in the 21st century consist of 16 skills that can be grouped into three categories, namely basic literacy, competence, and character quality. Basic literacy represents how students apply core skills to everyday tasks (Hikmawati, 2021). Literacy ability is a fundamental thing that must be possessed by students in facing the global era to be able to meet the needs of life in various situations. Literacy is the ability to read and write (Alfin, 2018; Christison & E. Murray, 2020). Literacy success requires the development of decoding skills, vocabulary, fluency, understanding, and the development of learning strategies to construct meaning from texts (Lysenko et al., 2019). Literacy and numeracy skills are essential for students because these two literacy skills are key in facing challenges in the global era (Muliani et al., 2021).

Numerical literacy is the knowledge, abilities, and skills possessed by students related to basic mathematics in the use of various numbers, and symbols, analyzing information in various forms of graphs, tables, charts, and so on to solve practical problems in the context of everyday life (Mahmud & Pratiwi, 2019; Megawati & Sutarto, 2021; I. L. Sari et al., 2021). The scope of numeracy is the ability to apply mathematical concepts and rules in real everyday situations when the problems are often unstructured (Mellyzar et al., 2022; Novita et al., 2022). In general, numeracy activities have a positive relationship with children's numeracy skill (Cheung et al., 2021). Important numeracy skills are mastered to reduce the risk of students' difficulties in mathematics (Dessemontet et al., 2020; Hellstrand et al., 2020). Mathematics as a path or counting element is used in

How to Cite:
Muliani, M., Novita, N., Mellyzar, M., Pasaribu, A. I., & Fadli, M.R. (2022). Analysis of the Characteristics of the Ethnoscience-Based Numeracy Test Instrument Using the Rasch Model. *Jurnal Penelitian Pendidikan IPA*, *8*(5), 2176–2183. https://doi.org/10.29303/jppipa.v8i5.2285

subjects in developing students' critical thinking skills to solve problems (Sellars, 2017).

Teaching mathematics in schools aims to equip students with mathematical literacy and the ability to use and apply mathematical knowledge in real-life situations that occur outside of school (Abdullah & Suhartini, 2017; Sumirattana et al., 2017). The results of a research study in England revealed that people with low math skills have low salaries. About 50% of men with low math skills have low incomes because math skills are needed in society, math skills are as important as reading and writing skills (Desoete & Baten, 2022).

Learning numeracy literacy can be developed by combining numeracy with the local cultural context, context in learning can be in the form of the context of learning motivation, national character, attitudes, knowledge, and or scientific skills (Indrawan & I Gede Jaka, 2021; Pinxteren et al., 2021). Learning that pays attention to local cultural wisdom or ethnoscience is one of the things that need to be considered in curriculum development in Indonesia. Ethnoscience is knowledge acquired by one's language and culture that can be verified and this can be innovated in science-based learning in the classroom (Imansari et al., 2018). Puspasari et al (2019) stated that Ethnoscience is meaningful learning that allows students to learn while doing or "*learning by doing*" which allows students to be able to connect the learning materials studied with the context of everyday life.

Based on the results of observations made in Lhokseumawe City Junior High Schools, namely SMPN Arun and SMPS Sukma Bangsa, it was found that the numeracy ability score was in the low category with an average student test score of 46.9 (out of 149 students). In addition, if you look at the results of the 2018 PISA results, Indonesia is always in the bottom position, is ranked 73rd out of 79 countries in the category of low numeracy literacy ability with an average math score of 379 (Hewi & Shaleh, 2020). This low numeracy literacy ability is also supported by research conducted by (R. H. N. Sari & Ariyadi, 2017) which states that students' mathematical literacy on understanding indicators is in a low category and other process indicators are in the very low category. (Winata et al., 2021) stated that the numeracy skills of class XI MA Darul Ma'wa Plandirejo students, Plumpang sub-district, Tuban district were still low. It is known that 61.90% of students scored below 50. The percentage of incorrect answers for the 3 indicators was 64.76%; 48.57% and 44.67%. (Napsiyah et al., 2022) stated that students' mathematical numeracy skills were still low. This is evidenced by the average score of student test results that have been given with an average value of 51.14 (out of 19 students), there are still many students who get scores below the average. (Dedi et al., 2021) stated that students' numeracy literacy skills at the stage of understanding the problem 16% included in the very poor category, at the planning stage 8% included in the very poor category, at the stage of implementing the plan 27% included in the very poor category, while at the stage of reviewing 0 %, including the category of very less. So, these results indicate that the ability of students to solve PISA-type questions is still relatively low.

The problems in this study were reviewed based on observations in the field, PISA results from 2000 to 2018, and several studies that stated the low numeracy skills of Indonesian students. To overcome low numeracy skills, it is necessary to strengthen numeracy skills. The strengthening of numeracy carried out in this study was through the development of an ethnoscience-based numeracy test instrument. The purpose of this study was to analyze the characteristics of the ethnoscience-based numeracy test instrument using the *Rasch*. The Rasch model is used to prove that the instrument has a high level of validity and reliability. This is because the use of the Rasch model is a solution to the validity problem where the Rasch model provides useful statistics and offers convenience to investigate validity and reliability (Krisanda & Harjito, 2021; Napitupulu, 2017).

Method

This study focuses on analyzing ethnoscience-based numeracy test instruments using the Rasch model. This type of research is quantitative. The subjects of this study were students of SMPN Arun and SMPS Sukma Bangsa, located in Lhokseumawe City, totaling 123 students. The number of questions to be analyzed as a test instrument is 68 items. The questions used refer to the minimum competency assessment (AKM). The difficulty level of the questions starts from very easy, easy, difficult, and very difficult. The test results in the form of scores were analyzed using Winsteps software. From the output of the Winsteps software, several parameter items are obtained that fit the Rasch model. Parameters are seen in the form of validity, reliability, and level of difficulty of the test. Cronbach's alpha value is the result of the overall item reliability test (Azizah & Sapti, 2020). The logit value on the item measure is the result of the item difficulty level test. Meanwhile, Outfit MNSQ, Outfit Z STD, and the correlation value of the item with the question as a whole show the limit of items that are declared fit with the model. That is if the Outfit MNSQ value is between 0.5 to 1.5; the Outfit ZSTD value is between -2.0 to 2.0, and the item correlation value with a total score is between 0.4 to 0.85 to measure the results of item validity.

Result and Discussion

This research was conducted to find data and information that can be used in analyzing the quality of the ethnoscience-based numeration test instrument empirically using the Rasch model. The Rasch model aims to develop objective measurements, by applying the right basic principles including 1) carrying out an appropriate estimation process; 2) giving a linear measure; 3) resolving lost data; 4) producing measurement instruments that are independent of the parameters studied; 5) find data that is not appropriate (misfits) or not common (outliers) (Khine, 2020). Analysis of the quality of test instruments is an activity carried out to examine each item of the instrument by collecting information from student answers to obtain a quality instrument (Ayub et al., 2020). The analysis of ethnoscience-based numeration instruments is seen from several aspects including reliability, validity, and the level of difficulty of the instrument, with the help of the Rasch model and Winsteps software. The data analyzed in this study is data from the test results of an ethnoscience-based numeration instrument consisting of 68 multiple-choice questions that have been tested on 123 junior high school students in the city of Lhokseumawe. The research data analyzed using Rasch modeling in detail are as follows:

Instrument Validity

The first analysis performed on the test instrument is constructed validity analysis. In the analysis using the Rasch model, the interpretation of construct validity and content validity measurements can be evaluated more precisely. Instrument validity is used to test whether ethnoscience-based numeracy instruments can be used to measure students' numeracy skills. The quality of the items from the validity aspect can be determined if they meet the following criteria (Sumintono & Widhiarso, 2015) (a)The outfit MNSQ (Mean Square) value received is: 0.5 < Outfit – MNSQ < 1.5 (b) outfit ZSTD Value (Z– The accepted standards are: –2.0 < ZSTD < +2.0 (c) measure Corr (Point Measure Correlation): 0.4 < Point Measure Corr < 0.85

Based on the results of the measure order analysis, it is known that there are four questions in the "maximum measure", namely questions number 14, 42, 67, and 68 which means that these four questions cannot be maintained because none of the students can answer the questions because they are too difficult. The questions that can still be defended for re-examination are 64 questions. The results of the analysis showed that all items that met the MNSQ value obtained a value of 0.5-1.5 with good criteria for measurement. Based on the ZSTD value analysis shown in the table, instrument item number 8, 9,11, and 12 did not meet the criteria.

Table 1. Results of ZSTD value analysis (Sumintono & Widhiarso, 2015)

Value ZSTD	Implications for Measurement	No Item
≥ 3.0	Data is not expected if it fits the model	8
2.0 - 2.9	Data seems unpredictable	9, 11, 12
-1 ,9 - 1.9	Data has logical approximations	1,2,3,4,5,6,7,10,13,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30, 32,32,33,34,35,36,37,38,39,40,41,43,44,45,46,50,53,55,56,57,58,59, 62, 63,64,65,66 -2
≤ 2.0	Data is too predictable	47, 48, 49, 51, 52,54, 60, 61

Table 2. The results of the analysis of the Pt-Measure Corr (Alagumalai et al., 2005)

Pt-Measure Corr Value	Measurement Criteria	No Item
> 4	Very Good	32, 34, 40, 46, 47, 48, 49, 50, 51 , 52, 53, 54, 55, 56, 59, 60, 61,
0.3-0.39	Good	31, 38, 41, 44, 62, 65
0.20-0.29	Enough	10, 17, 18, 20, 21, 23, 30, 35, 39
0.00-0.19	Unable to discriminate against	1, 2, 3,4,5, 6, 7, 13, 15, 16,19, 22, 24, 25, 26 , 28, 29, 33, 36, 37, 43,45, 57, 58, 63, 64, 66
< 0,00	Require re-examination	8, 27

Pt-Measure Corr data analysis shows that two instrument items require re-examination, namely 8 and 27. If the item meets the three criteria, it can be said that the item is "Valid" and can be used for sure, whereas if the item only meets two criteria or one criterion, the item can still be maintained with slight revisions so that it is still categorized as " valid" and can be used after being revised, but if it does not meet the three criteria, it can be said The items are categorized as "invalid" so they cannot be maintained and used. The results of the item fit analysis include the MNSQ value, ZSTD value, and the Pt-Measure Core value, obtained 9 item instrument questions with numbers 8, 9, 11, 12, 14, 27, 42, 67, and 68 invalids and cannot be maintained. Valid and usable test instruments are 59 questions with details of instruments with valid criteria totaling 32 questions with numbers 10, 17, 18, 20, 21, 23,30, 31, 32, 34, 35, 38, 39, 40, 41 , 44, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 59, 60, 61,62, and 65. While 27 consists of numbers 1, 2, 3,4, 5, 6, 7, 13, 15, 16,19, 22, 24, 25, 26, 28, 29, 33, 36, 37, 43,45, 57, 58, 63, 64, and

66 need to be revised for improvement before being published and used.

Instrument

Reliability Reliability is a value that indicates an instrument is reliable or trusted so that even if it is used many times it will produce the same or consistent results. The value of the reliability coefficient is expressed in numbers that indicate the level of reliability empirically. A reliable measuring instrument if the correlation coefficient between the results of measuring two parallel tests is high, then the consistency between the two is getting better. The interpretation of Item Reliability and Person Reliability value criteria is shown in Table 3.

Table 3. Reliability Criteria in Rasch Modeling

Reliability Value (Person/Item)	Interpretation
> 0.94	Special
0.91-0.94	Very
0.81-0.90	Good
0.67-0.80	Enough
< 0.67	Weak

Table 4. Reliability Test Interpretation based on Cronbach Alpha

Value Cronbach Alpha	Interpretation
> 0.8	Very Good
0.7 < < 0.8	Good
0.6 < < 0.7	Enough
0.5 < < 0.6	Poor
< 0.5	Very Poor

The reliability of respondents' answers can be seen from the output Summary Statistics presented in the Figure 1.

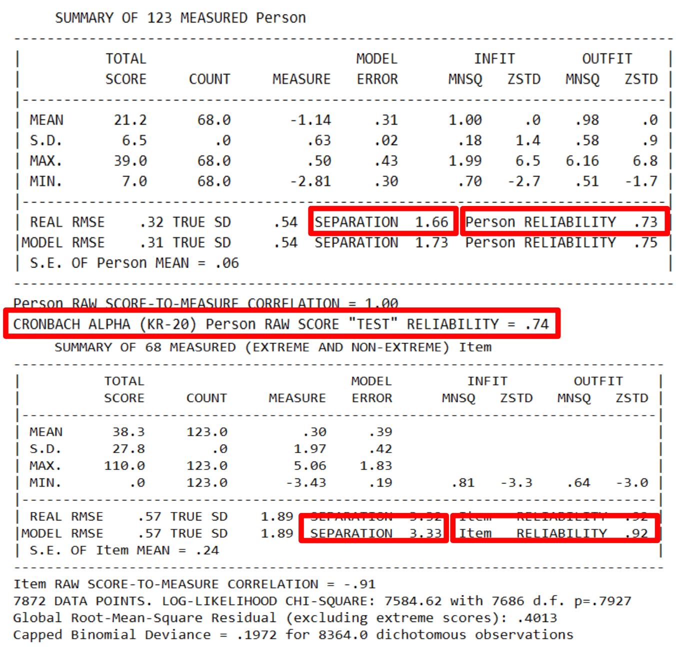
```
     SUMMARY OF 123 MEASURED Person
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      21.2      68.0      -1.14     .31      1.00     .0     .98     .0 |
| S.D.       6.5        .0        .63     .02       .18    1.4     .58     .9 |
| MAX.      39.0      68.0        .50     .43      1.99    6.5    6.16    6.8 |
| MIN.       7.0      68.0      -2.81     .30       .70   -2.7     .51   -1.7 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .32 TRUE SD    .54  SEPARATION  1.66  Person RELIABILITY  .73 |
|MODEL RMSE    .31 TRUE SD    .54  SEPARATION  1.73  Person RELIABILITY  .75 |
| S.E. OF Person MEAN = .06                                                   |
-------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .74
     SUMMARY OF 68 MEASURED (EXTREME AND NON-EXTREME) Item
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      38.3     123.0        .30     .39                                 |
| S.D.      27.8        .0       1.97     .42                                 |
| MAX.     110.0     123.0       5.06    1.83                                 |
| MIN.        .0     123.0      -3.43     .19       .81   -3.3     .64   -3.0 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .57 TRUE SD   1.89  SEPARATION  3.32  Item   RELIABILITY  .92 |
|MODEL RMSE    .57 TRUE SD   1.89  SEPARATION  3.33  Item   RELIABILITY  .92 |
| S.E. OF Item MEAN = .24                                                     |
-------------------------------------------------------------------------------
Item RAW SCORE-TO-MEASURE CORRELATION = -.91
7872 DATA POINTS. LOG-LIKELIHOOD CHI-SQUARE: 7584.62 with 7686 d.f. p=.7927
Global Root-Mean-Square Residual (excluding extreme scores): .4013
Capped Binomial Deviance = .1972 for 8364.0 dichotomous observations
```

Figure 1. Summary of Rasch Modeling Statistics

Summary statistics provide overall information about the quality of the instruments used, the quality of respondents, as well as the interaction between persons and items. Based on the data in Table 5. above, it can be explained that the reliability in the Rasch modeling analysis is divided into three parts, namely 1) respondent reliability, 2) item reliability, and 3) overall reliability. The Person Reliability value is 0.73 which is included in the "enough" category and the Item Reliability value is 0.92 in the "very good" category. While the overall reliability is shown by the Cronbach Alpha value which measures the reliability of the interaction between the person and the item about the instrument as a whole, it obtained a value of 0.74 with the "good" category.

Rasch modeling can also be used in analyzing the level of individual abilities to distinguish the abilities of students who can answer questions and those who are unable to answer questions. In addition, it can also be used to identify groups of respondents based on the respondent separation index. The greater the value of item separation, the better the quality of the instrument, because it can identify groups of respondents and groups of items. To find out the grouping more accurately, the strata Equation 1.

$$H = \frac{[(4xSeparation)+1}{3} \quad (1)$$

Based on the data in table 5, it is known that the instrument item separation value is 3.32, so the value of H=4.76 is rounded to 5, this indicates that There are five groups of questions. Meanwhile, for the respondents, a separation value of 1.66 was obtained with H=2.55 rounded up to 3, which indicates that there are three groups of respondents. So, this instrument can be used to measure respondents who have ethnoscience-based numeracy skills in the low medium and high categories.

Instrument Difficulty Level (Item Measure)

A test instrument is said to be feasible if it can reach all respondents' abilities and respondents can reach all instrument difficulty levels. To find out the level of difficulty of the test instrument, we can see through the logit value of the item measure. In Rasch modeling, the difficulty level of the items is categorized based on the Logit Measure and the Logit item Standard Deviation (SD) value and is divided into four categories as shown in Table 6 (Sumintono & Widhiarso, 2015).

Table 6. Criteria for Difficulty Level of Items with Rasch Modeling

Measure value (logit)	Interpretation of Difficulty Item Item
Measure logit > SD logit	Very difficult item
0≤Measure logit≤ SD logit	Difficult item
-SD Measure logit 0	Easy item
Measure logit < - SD logit	Very easy item

Based on the data analysis that has been done Information on the grouping of items based on the level of difficulty is obtained as shown in Table 7.

Table 7. Distribution of Item Difficulty Levels with Rasch Modeling

Measure Value (logit)	Interpretation of Item Difficulty	Number of Items
Measure logit > SD logit	Very difficult	13 items (14, 42, 67, 68, 25, 28, 37, 66 , 36, 29, 58, 63, and 2)
0≤Measure logit≤ SD logit	Difficult items	19 items (64, 26, 43, 56, 27, 45, 59, 15, 57, 46, 53, 13, 44, 65, 55, 1.54, 50, and 5)
-SD Measure logit 0	Easy items	33 items (51, 22, 24, 52, 61, 48, 49, 16, 60, 11, 12, 8, 62 , 30, 10, 41, 47, 9, 7, 33, 39, 35, 21, 20, 23, 6, 18, 34, 38, 40, 19, 32, and 31)
Measure logit < - SD logit	Item very easy	3 items (17, 3, and 4)

Data analysis of research results using the Winsteps program provides information both in terms of items and respondents to the ethnoscience-based numeration test instrument which is analyzed using Rasch modeling. Some of the characteristic components of an instrument being analyzed include validity, reliability, measuring the power of the instrument, instrument validity, level of difficulty, and the quality of the instrument in measuring several groups of subjects. Assessment instruments analyzed using Rasch modeling can discuss instrument assessments more specifically. Efforts to determine student competence in a material can be obtained by focusing on the discussion of the test instrument material. In this article, the author analyzes an ethnoscience-based numeracy test instrument.

The analysis of the validity of the ethnoscience-based numerical instrument in the Winsteps program is called the fit and misfit test (valid and invalid items), which can be done by analyzing the output of the item fit order. Item fit can explain whether the items function normally to take measurements or not. Outfit means-square, outfit z-standard, and point measure correlation are the criteria used to see the level of item fit (Boone et al., 2013). From all test instruments analyzed by Rasch modeling, it was found that there were 9 items with numbers 8, 9, 11, 12, 14, 27, 42, 67, and 68 that did not meet the three validity criteria including MNSQ outfit, ZSTD Outfit, and Point Measure Correlation (Pt Measure Corr), so it can be said that the item is invalid (Misfit) so it cannot be maintained and is not suitable for use. Meanwhile, for other items, because they meet at least one criterion to be said to be a fit or valid item, the value of other items that do not match the two existing criteria but meet one criterion can still be said to be a Fit or valid item and can still be maintained for use in research. Of the 68 instrument items, 59 items were obtained that were declared valid, with details of 32 valid items eligible for use and 27 valid items that needed to be revised for improvement before being published and used.

Reliability is the constancy or determination of a measuring instrument in measuring what is being assessed. A reliable instrument is if the measurement results using the instrument provide stable and relatively the same results (Sumintono & Widhiarso, 2015). Through Table 5, which is the Summary Statistics output for the ethnoscience-based numerical instruments in the results above, the following can be observed (1) Cronbach's alpha value (measuring reliability, namely the interaction between respondents and items) is 0.74 with the interpretation of the score including the criteria "Good". So, the instrument used can be said to be reliable. (2) The Person Reliability value in the output above is 0.73 which indicates that the consistency of the answers from the respondents is said to be "Enough". While the Item Reliability value of 0.92 indicates a "very good" value criterion so that it can be said that the quality of the items used in the instrument is reliable. (3) The grouping of people and items can be seen from the separation value. If the value of separation is greater, the quality of the instrument in terms of overall respondents and items is better, because it can identify groups of respondents and groups of items. Based on Table 5. The score of separation items is 3.32, and the value of H=4.76 is rounded to 5, this shows that there are five groups of items. As for the separation of persons, the score is 1.66 with H=2.55 rounded up to 3, which indicates that the group of persons can be divided into three groups of persons. So this instrument can be used to measure subjects who have ethnoscience-based numeracy skills in the low, medium, and high categories.

Based on the results of the analysis above, the interpretation of Cronbach Alpha is 'good'. This shows that there is a match between the item and the person (respondent). Then the consistency of answers from respondents (personal reliability) can be said to be 'enough' with the quality of the instrument items (item reliability) being 'very good. So it can be concluded that the researched ethnoscience-based numeration instrument can be said to be reliable.

The results of the analysis of the difficulty level of ethnoscience-based numerical instruments for each category are very difficult, there are 13 questions, 19 questions are difficult, 33 questions are easy, and 3 questions are very easy. From the data analysis, it was obtained 4 questions that were in the "maximum measure" namely questions number 14, 42, 67, and 68, which means that these four questions cannot be maintained because they are too difficult and none of the students can answer the question. Different levels of item difficulty are required to identify different abilities. In very difficult category questions, the percentage of respondents who answered correctly tended to be the lowest compared to other question categories. However, the categories of questions are difficult, easy, and very easy, and the percentage of students who answer correctly tends to vary with several possibilities, such as respondents answering questions by guessing. This possibility can occur in several questions that are classified as: difficult, easy, and very easy because of the pattern of presentation of the questions or differences in the level of ability of the respondents. Overall, from the results of the research data analysis, it was found that the ethnoscience-based numeration instrument in this study was declared "reliable" and 59 of the 68 items tested were declared "valid".

Conclusion

The use of the Rasch model in analyzing the characteristics of ethnoscience-based numerical instruments produces specific and holistic information about the instrument under study. Based on the results of the analysis, it was found that the ethnoscience-based numeration instrument in this study was declared reliable. 59 of the 68 items were said to be valid, with details of 32 valid items eligible to be used and 27 valid items needing to be revised for improvement before being used. This instrument can be used to measure subjects who have ethnoscience-based numeracy skills in the low, medium, and high categories, so to examine ethnoscience-based numeracy skills students can use this instrument.

Acknowledgements

The research team would like to thank the Universitas Malikussaleh Research and Community Service Institute (LPPM) for entrusting the research team to provide internal non-tax revenues so that this research can be carried out.

References

Abdullah, A. A., & Suhartini, S. (2017). Meningkatkan Kemampuan Berpikir Kritis Melalui Pembelajaran Statistika Berbasis Pendidikan Politik Di Lingkungan Sekolah. *Jurnal Gantang*, *2*(1), 1–9. https://doi.org/10.31629/jg.v2i1.63

Alagumalai, S., Curtis, D. D., & Hungi, N. (2005). *Applied Rasch measurement: A book of exemplars*. Dordrecht, The Netherlands: Springer.

Alfin, J. (2018). Membangun Budaya Literasi dalam Pembelajaran Bahasa Indonesia Menghadapi Era Revolusi Industri 4.0. *Pentas: Jurnal Ilmiah Pendidikan Bahasa Dan Sastra Indonesia*, *4*(2), 60–66.

Ayub, M. R. S. S. N., Istiyono, E., Munadi, S., Permadi, C., Pattiserlihun, A., & Sudjito, D. N. (2020). Analisa Penilaian Soal Fisika Menggunakan Model Rasch Dengan Program R. *Jurnal Sains Dan Edukasi Sains*, *3*(2), 46–52. https://doi.org/10.24246/juses.v3i2p46-52

Azizah, A., & Sapti, W. (2020). Penggunaan Model Rasch Untuk Analisis Instrumen Tes Pada Mata Kuliah Matematika Aktuaria. *Jurnal Pendidikan Matematika*, *3*(1), 45–50. https://doi.org/10.30598/jupitekvol3iss1pp45-50

Boone, W. J., Staver, J. R., & Yale, M. S. (2013). *Rasch analysis in the human sciences*. Springer Science & Business Media.

Cheung, S. K., Katrina May, D., Xiujie, Y., Mohseni, F., & Catherine, M. (2021). Home Literacy and Numeracy Environments in Asia. *Frontiers in Psychology*, *12*, 1–20. https://doi.org/10.3389/fpsyg.2021.578764

Christison, M., & E. Murray, D. (2020). An overview of multilingual learners' literacy needs for the 21st century. In *Handbook of research on cultivating literacy in diverse and multilingual classrooms* (pp. 1–20). https://doi.org/10.4018/978-1-7998-2722-1

Dedi, D., Dekriati, A., & Samuel, R. M. M. (2021). Analisis Kemampuan Literasi Numerasi Siswa SMP Negeri 2 Wanukaka. *… Matematika Sumba*, *3*(2), 125–132. https://doi.org/https://doi.org/10.53395/jppms.v3i2

Desoete, A., & Baten, E. (2022). Math Learning in Grade-4 and 5: What Can We Learn Form The Opportunity-Propensity Model? *International Electronic Journal of Elementary Education*, *14*(3), 213–225. https://doi.org/10.26822/iejee.2022.240

Dessemontet, R. S., Opitz, E. M., & Schnepel, S. (2020). The profiles and patterns of progress in numerical skills of elementary school students with mild and moderate intellectual disability. *International Journal of Disability, Development and Education*, *67*(4), 409–423. https://doi.org/10.1080/1034912X.2019.1608915

Hellstrand, H., Korhonen, J., Räsänen, P., Linnanmäki, K., & Aunio, P. (2020). Reliability and validity evidence of the early numeracy test for identifying children at risk for mathematical learning difficulties. *International Journal of Educational Research*, *102*, 1–10. https://doi.org/10.1016/j.ijer.2020.101580

Hewi, L., & Shaleh, M. (2020). Refleksi Hasil PISA (The Programme For International Student Assesment): Upaya Perbaikan Bertumpu Pada Pendidikan Anak Usia Dini). *Jurnal Golden Age*, *4*(01), 30–41. https://doi.org/10.29408/jga.v4i01.2018

Hikmawati, H. (2021). Kegiatan Analisis Artikel Tentang Etnosains Dan Kearifan Lokal Masyarakat Suku Sasak Untuk Mengembangkan Literasi Sains Dan Literasi Budaya Mahasiswa. *Jurnal Pendidikan Dan Pengabdian Masyarakat*, *4*(3), 331–339. https://doi.org/10.29303/jppm.v4i3.2859

Imansari, M., Sudarmin, & Sumarni, W. (2018). Analisis Literasi Kimia Peserta Didik Melalui Pembelajaran Inkuiri Terbimbing Bermuatan Etnosains. *Jurnal Inovasi Pendidikan Kimia*, *12*(2), 2201–2211. https://doi.org/10.15294/jipk.v12i2.15480

Indrawan, I. P. O., & I Gede Jaka, M. (2021). E-Learning Terintegrasi Kearifan Lokal Bali Berbasis 4c pada Mata Pelajaran IPA. *Jurnal Pedagogi Dan Pembelajaran*, *4*(3), 511–521. https://doi.org/10.23887/jp2.v4i3.41415

Karatas, K., & Arpaci, I. (2021). The role of self-directed learning, metacognition, and 21st century skills predicting the readiness for online learning. *Contemporary Educational Technology*, *13*(3), 1–13. https://doi.org/10.30935/cedtech/10786

Khine, M. S. (2020). *Rasch measurement*. .Springer Singapore. https://doi.org/10.1007/978-981-15-1800-3

Krisanda, Y. I., & Harjito, H. (2021). PENGGUNAAN RASCHMODEL UNTUK ANALISIS INSTRUMEN TES LITERASI. *Journal of Chemistry In Education*, *2*(3), 55–61. https://doi.org/10.1111/j.1949-8594.1902.tb00418.x

Lysenko, L., Abrami, P. C., Wade, C. A., Marsh, J. P., WaGioko, M., & Kiforo, E. (2019). Promoting young Kenyans' growth in literacy with educational technology: A tale of two years of implementation. *International Journal of Educational Research*, *95*, 176–189. https://doi.org/10.1016/j.ijer.2019.02.013

Mahmud, M. R., & Pratiwi, I. M. (2019). Literasi numerasi siswa dalam pemecahan masalah tidak terstruktur. *Kalamatika: Jurnal Pendidikan Matematika*, *4*(1), 69–88. https://doi.org/10.22236/KALAMATIKA.vol4no1.2019pp69-88

Megawati, L. A., & Sutarto, H. (2021). Analysis numeracy literacy skills in terms of standardized math problem on a minimum competency assessment. *Unnes Journal of Mathematics Education*, *10*(2), 155–165. https://doi.org/10.15294/ujme.v10i2.49540

Mellyzar, M., Unaida, R., Muliani, M., & Novita, N. (2022). Hubungan Self-efficacy dan Kemampuan Literasi Numerasi Siswa: Ditinjau Berdasarkan Gender. *Lantanida Journal*, *9*(2), 104–113. https://doi.org/10.22373/lj.v9i2.11176

Muliani, M., Marhami, M., & Lukman, I. R. (2021). Persepsi Mahasiswa Calon Guru Tentang Literasi Sains. *JISIP (Jurnal Ilmu Sosial Dan Pendidikan)*, *5*(1). https://doi.org/10.36312/jisip.v5i1.1575

Napitupulu, D. (2017). Analisis model rasch pada instrumen keberterimaan marka optik penanda jarak aman kendaraan. *Jurnal Penelitian Transportasi Darat*, *19*(1), 37–48. https://doi.org/10.25104/jptd.v19i1.604

Napsiyah, N., Nurmaningsih, N., & Rahman, H. (2022). Analisis Kemampuan Numerasi Matematis Siswa Berdasarkan Level Kognitif pada Materi Kubus dan Balok. *JagoMIPA: Jurnal Pendidikan Matematika Dan IPA*, *2*(2), 45–59. https://doi.org/10.53299/jagomipa.v2i2.183

Novita, N., Muliani, M., & Mellyzar, M. (2022). Pelatihan Pengembangan Soal Matematika dan Sains Berbasis numerasi Pada Guru Untuk Menunjang Asesmen Nasional. *SELAPARANG Jurnal Pengabdian Masyarakat Berkemajuan*, *6*(1), 486–493. https://doi.org/10.31764/jpmb.v6i1.7761

Pinxteren, B. van, Emirhafizović, M., Dailidienė, I., Figurek, A., Zamfir, R. H.-C., & Mętrak, M. (2021). Education for resilience: How a combination of systemic and bottom-up changes in educational services can empower dryland communities in Africa and Central Asia. *Journal of Asian and African Studies*, *56*(6), 1271–1285. https://doi.org/10.1177/00219096209601

Pratiwi, S. N., Cari, C., & Aminah, N. S. (2019). Pembelajaran IPA abad 21 dengan literasi sains siswa. *Jurnal Materi Dan Pembelajaran Fisika*, *9*(1), 34–42. https://doi.org/10.20961/jmpf.v9i1.31612

Puspasari, A., Indah, S., Kurniawati, L., Resiana Ridha, U., Indra, G., & Ika Candra, S. (2019). Implementasi Etnosains dalam Pembelajaran IPA di SD Muhammadiyah Alam Surya Mentari Surakarta. *SEJ (Science Education Journal)*, *3*(1), 25–31. https://doi.org/10.21070/sej.v3i1.2426

Sari, I. L., Irawan, E., Aristiawan, A., & Rokmana, A. W. (2021). Analisis Tingkat Penalaran Peserta Didik SMP dalam Memecahkan Masalah Soal Evaluasi Berbasis Literasi Numerasi. *Jurnal Tadris IPA Indonesia*, *1*(3), 333–342. https://doi.org/10.21154/jtii.v1i3.135

Sari, R. H. N., & Ariyadi, W. (2017). Mathematical literacy of senior high school students in Yogyakarta. *Jurnal Riset Pendidikan Matematika*, *4*(1), 100–107. https://doi.org/10.21831/jrpm.v4i1.10649

Sellars, M. (2017). Numeracy across the curriculum - A pathway to critical thinking. *International Journal of Innovation, Creativity and Change*, *3*(2), 75–83.

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan rasch pada assessment pendidikan*. Trim komunikata.

Sumirattana, S., Makanong, A., & Thipkong, S. (2017). Using realistic mathematics education and the DAPIC problem-solving process to enhance secondary school students' mathematical literacy. *Kasetsart Journal of Social Sciences*, *38*(3), 307–315. https://doi.org/10.1016/j.kjss.2016.06.001

Widyastika, D., Sitorus, R. H., & Lubis, S. J. (2022). Literasi Sains dan Pendidikan Karakter pada Pembelajaran IPA Abad 21. *Journal on Teacher Education*, *3*(3), 302–309. https://doi.org/10.31004/jote.v3i3.4922

Winata, A., Ifa Seftia Rakhma, W., & Sri Cacik. (2021). Analisis Kemampuan Numerasi dalam Pengembangan Soal Asesmen Kemampuan Minimal pada Siswa Kelas XI SMA untuk Menyelesaikan Permasalahan Science. *Jurnal Educatio FKIP UNMA*, *7*(2), 498–508. https://doi.org/10.31949/educatio.v7i2.1090