Jurnal Penelitian dan Evaluasi Pendidikan
Volume 28, No. 1, June 2025 (1-12)
Online: http://journal.uny.ac.id/index.php/jpep

Validity and reliability of students' mathematical communication instruments in class vii middle school statistics material

Syahfira Audiwinanda; Ali Mahmudi
Universitas Negeri Yogyakarta, Indonesia
*Corresponding Author. E-mail: syahfiraaudiwinanda.2022@student.uny.ac.id

ARTICLE INFO

Article History
Submitted:
04 February 2024
Revised:
20 May 2024
Accepted:
18 August 2024

Keywords
instrument; reliability; validit; mathematical communication

Scan Me:

ABSTRACT

This research aims to reveal content validity, construct validity, and reliability of mathematical communication ability instruments. This research involved 100 junior high school students. The instrument has been validated, contain 12 multiple-choice questions and three describing mathematical communication skills, which use the V Aiken index to measure content validity. Meanwhile, construct validity was proved using factor analysis using the Kaiser Meyer-Olkin (KMO) test, and Cronbach's Alpha test for reliability estimation by R Studio. The level of difficulty, different strengths, alternative answers, and final conclusions are calculated using Anbuso version 8.0. The research results show that 1) The instrument meets the content validity criteria from the material, construction and linguistic aspects based on the satisfaction of 4 expert validators, proven by calculating the V Aiken index for all valid question items. 2) All items meet the construct validity criteria. 3) The estimated reliability for all types of questions is 0.812 for multiple choice questions with an SEM of 2.5 and 0.868 for essay questions with an SEM of 2.1, so the instrument is reliable overall. 4) The distinguishing power of the good category is 87%, while the quite good category is 13%. 5) The difficulty level for the difficult category is 7%, and the medium category is 93%. 6) All alternative answers work well. 7) Final conclusion: 93% of the 16 questions on the mathematical communication ability instrument are acceptable, and 7% need to be changed. 7% (1 question) are questions with indicators connecting diagrams, graphs, tables to mathematical idea.

To cite this article (in APA style):
Audiwinanda, S., & Mahmud, A. (2025). Validity and reliability of students' mathematical communication instruments in class vii middle school statistics material. *Jurnal Penelitian dan Evaluasi Pendidikan, 29(1)*, 1-12 doi: https://doi.org/10.21831/pep.v29i1.71234

INTRODUCTION

One of the goals of learning mathematics is to equip students to improve mathematical communication. This is stated in BSKAP 08 of 2022 discussing Learning Achievements, stating that one of the objectives of learning mathematics is to communicate ideas using symbols, tables, diagrams or other tools to clarify concepts, situations or problems and represent them in the form of symbols or mathematical models. In addition, NCTM (2000) sets school mathematics standards that students must be able to communicate mathematically. It means that it is very important for students to be able to communicate mathematically. Mathematical communication is a way for students to express their mathematical ideas verbally, in writing, drawings, diagrams, using objects, presenting them in algebraic form, or using mathematical symbols (Qohar & Fazira, 2022). According to Nurlaila, et al. (2018) stated that mathematical communication is a way to express a mathematical idea orally or in writing in the form of pictures, algebra or diagrams. Communication skills can also be defined as the ability to understand and accept other people's mathematical ideas or thoughts carefully, analytically, critically and evaluatively to sharpen understanding (Babys, 2020). Furthermore K et al. (2021) that mathematical communication skills include students' skills in conveying their understanding orally and in writing using mathematical language.

10.21831/pep.v29i1.71234

Communication is important in mathematics and mathematics education because it allows students to share ideas and understand what others are saying. Mathematical communication has an important role in gaining mathematical knowledge and expressing mathematical ideas (Kuswandi & Astuti, 2019). However, based on the research results of Wijayanto et al. (2023), junior high school students' mathematical communication skills are still relatively low. In line with the research results of Andini and Marlina (2021) that, the mathematical communication skills of junior high school students are still relatively low. Therefore, ways are needed to improve students' mathematical communication skills.

So that students can improve their mathematical communication skills, they need to often practice working on mathematical ability questions. Therefore, a good instrument is needed to measure mathematical communication skills. An instrument is said to be good if it can measure the level of effectiveness of a measuring instrument and measure the extent to which the measuring instrument can be trusted (Alfiatunnisa et al., 2022). A good instrument needs to be a valid and reliable instrument. If the assessment instrument is invalid and unreliable, the test results are doubtful. Assessment tools must have the right quality to measure student competency accurately (Sa'idah et al., 2018). The results of proving validity, reliability estimates, level of difficulty, differentiability, and how effective alternative answers are can determine the quality of an instrument (Friatma & Anhar, 2019; Jannah et al., 2021).

Revita et al. (2018) based on the results of their research on the analysis of mathematical communication ability instruments on function and relationship material, there were 4 questions with a medium level of difficulty and two questions with an easy level of difficulty. There are 5 questions with good different powers and one with quite good different powers. So, there is still a need to improve the question-making process to increase the warehouse of good questions, especially in other mathematical materials, such as statistics. Statistics material is important because it explains understanding the basic concepts of data presentation techniques and interpreting the meaning of the data presented (Khadijah et al., 2018).

Several previous studies have developed mathematical communication instruments, including Alamsyah (2015); this research developed mathematical communication instruments for junior high school students but only focused on the material on Systems of Linear Equations in One Variable. Then Azmi and Salam (2020) also developed a mathematical communication ability instrument which focused only on rectangular material. Wahyuni (2022) also analyzed the validity and reliability of mathematical communication ability instruments only on relations and functions. However, there has been no development of instruments for mathematical communication skills in statistical material, where statistical material is closely related to mathematical communication. In statistics material, there is much discussion about the presentation and interpretation of data in the form of pictures, diagrams and graphs. Apart from that, it is hoped that the developed instrument can streamline processing time, improve students' ability to choose the correct answer and to determine the extent of students' written mathematical communication skills, the researchers created two types of questions, namely multiple choice questions and essays. Based on the above, the main aim of this research is to test the validity and reliability of the instrument for students' mathematical communication skills in statistics material for class VIII SMP.

RESEARCH METHOD

This research is part of research into the development of mathematical communication ability instruments. Instruments that meet validity standards and reliability standards can be used for the measurement stage. The overall research stages presented in Figure 1 refer to the creation of test instruments created by (Supahar & Prasetyo, 2015).

10.21831/pep.v29i1.71234

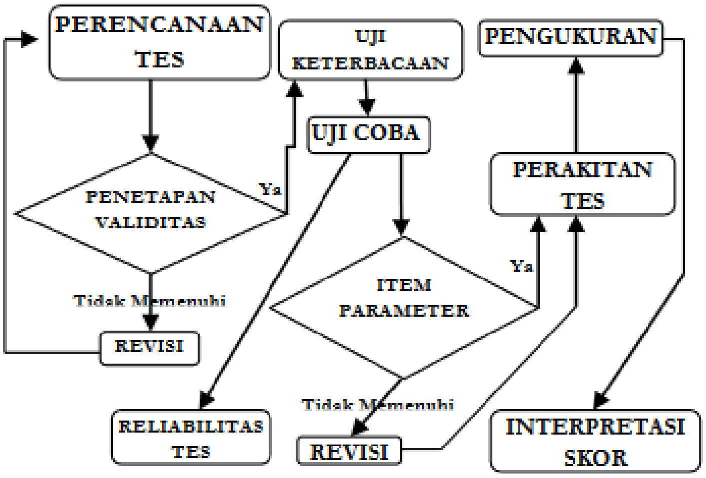

Figure 1. Scheme of Instrument Development Research Procedures Mathematical Communication

Mathematical communication capability instruments were developed based on statistical material and indicators of mathematical communication, namely connecting diagrams, graphs, and tables into mathematical ideas, using formulas correctly, and explaining the reasons for solving the problem in their own language. The data obtained to determine the validity and reliability of the instrument is obtained in several ways. The instrument's content validity is obtained by giving questionnaires to experts to test content validity.

Empirical validation and reliability were obtained through instrument testing. The communication skills instrument tested was 15 questions, including 12 multiple choice questions and three essays, for 100 students in Class VIII of SMP Negeri 2 Gamping. Based on the results of students' answers, an analysis of validity verification (content and construct), reliability estimation, level of difficulty, distinguishing power and alternative answers was carried out. In validating the content, it is based on input from 4 expert validators and also proven using the V Aiken index calculation with the following formula 1.

$$V = \frac{\sum s}{n(e-1)} \quad (1)$$

V represents Aiken's validity index, which is used to measure the degree of agreement among validators regarding the relevance of an item. The value of s is obtained by subtracting the lowest score on the scale (l_0) from the score given by the validator (r). The symbol r refers to the rating assigned by a validator, while l_0 indicates the lowest possible score on the scale. n denotes the number of validators involved in the assessment, and e represents the number of response categories available on the rating scale. If the v aiken value is greater than the V table of 0.8800, then the question item is said to be valid. Meanwhile, to prove construct validity using Factor Analysis in the Kaiser Meyer-Olkin (KMO) test and reliability estimation using the Cronbach's Alpha test with the help of R Studio. For the level of difficulty, differentiation, alternative answers and final conclusions with the help of Anbuso version 8.0 based on Microsoft Excel developed by Ali Muhson.

Proving the validity of an instrument is carried out to determine its accuracy in measuring what it should measure. A tool is said to be sufficient for sampling if the KMO value is more than 0.5 and the significant value is less than 0.05 so that it can be analyzed using factor analysis.

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.71234

The instrument is said to be valid if the loading factor value is more than 0.5(Alfiatunnisa et al., 2022).

The reliability estimate for this question is calculated using the Cronbach's Alpha formula. According to (Istiyono, 2020), stated that if the Cronbach's Alpha value of an instrument is more than 0.6 then it is said to be reliable. Differential power is determined to show that students with above and below abilities have differences. Difference between each question item can be calculated using the following formula 2 (Boopathiraj & Chellamani, 2013):

$$DB = \frac{(WL - WH)}{n} \quad (2)$$

DB represents the discrimination power of a test item, which indicates how well the item distinguishes between high-performing and low-performing test takers. WL refers to the number of participants in the lower group who answered the item incorrectly, while WH refers to the number of participants in the upper group who also answered incorrectly. The variable n represents 27% of the total number of test takers, commonly used to define the size of the upper and lower groups in item analysis. The differential power coefficient of an item is said to be good if it exceeds 0.3. On the other hand, a coefficient that is in the range of 0.20 to 0.29 is considered quite good and if it is below 0.2, it is discarded (Magdalena et al., 2021). Table 1 displays the different power criteria used.

Table 1. Different Power Criteria for Question Items

Different power	Criteria
$DB < 0.2$	Bad
$0,2 \leq DB \leq 0.3$	Enough
$DB > 0.3$	Good

Calculating the difficulty level of test items is a way to calculate how difficult each test item is. The difficulty index shows how difficult or easy a task is. Formula 3 to determine the level of difficulty of each question item (Kholis, 2018):

$$D = \frac{Ba + Bb}{Ja + Jb} = \frac{B}{JS} \quad (2)$$

The difficulty index (D) indicates how challenging a question is for test-takers. To calculate it, we look at Ba, which is the number of correct answers from the upper group, and Bb, the number of incorrect answers from the lower group. These values are considered in relation to Ja and Jb, representing the total number of answer sheets from the upper and lower groups, respectively. In a broader sense, B refers to the total number of students who answered the question correctly across all groups, while JS stands for the total number of students who took the test. Together, these values help determine how accessible or difficult a particular item is for the test population. The criteria for a good level of difficulty is 0.3–0.7 (Solichin, 2017). Therefore, the difficulty level criteria are presented in Table 2.

Table 2. Criteria for Difficulty Level of Question Items

Difficulty Level	Criteria
$DP < 0.3$	Easy
$0,3 \leq DP \leq 0.7$	Currently
$DP > 0.7$	Difficult

10.21831/pep.v29i1.71234

According to the criteria, alternative answers are considered ineffective if they are answered by at least 5% of examinees. Determination of final conclusions in this research uses the following criteria. (1) Good if the discrimination power is good or quite good, the level of difficulty is moderate, and all alternative answers are effective; (2) if the discrimination power is good or quite large and the level of difficulty is moderate, but the alternative answer is not effective, correct the alternative; (3) good enough if the power difference is large or quite large but the difficulty level is easy or hard; and not Good if power difference is not good.

FINDINGS AND DISCUSSION

The mathematical communication ability instrument studied in this research results from developing mathematical communication questions in class VIII junior high school statistics material. In the initial stage, mathematical communication instruments were analyzed by experts from Yogyakarta State University Mathematics Education by paying attention to material, construction and language aspects. This content validation activity is carried out by providing validation sheets to obtain suggestions and input regarding the mathematical communication ability instruments that will be tested on students. These suggestions and input are used as a reference for revising the questions and stating that the communication skills instrument is valid.

Next, the researcher calculated content validity based on the assessment results of 4 experts using the V Aiken formula. Below is Table 3. Results of calculating the validity of the instrument content using the V Aiken formula.

Table 3. Instrument content validity using V Aiken

Indicators of Mathematical Communication Skills	Question Form	Question Number	V count	V Table	Criteria
Connecting diagrams, graphs, tables into mathematical ideas	Multiple choice	1	0.9375	0.8800	Valid
		4	0.9375	0.8800	Valid
		9	0.9375	0.8800	Valid
		10	0.9375	0.8800	Valid
Use the formula correctly	Multiple choice	2	0.9375	0.8800	Valid
		3	0.9375	0.8800	Valid
		5	0.9375	0.8800	Valid
		6	0.9375	0.8800	Valid
		7	0.9375	0.8800	Valid
		8	0.9375	0.8800	Valid
		11	0.9375	0.8800	Valid
		12	0.9375	0.8800	Valid
	Essay	13	0.9375	0.8800	Valid
Explain the reasons for solving the problem in your own language	Essay	14	0.9375	0.8800	Valid
		15	0.9375	0.8800	Valid

Proving the Validity of Mathematical Communication Ability Instruments

The results of proving the validity of the mathematical communication ability instrument on multiple choice questions with Factor Analysis on the Kaiser Meyer-Olkin (KMO) test using R studio are 0.91 > 0.5 for multiple choice questions dan 0.65 > 0.5 untuk soal essay. Based on the overall KMO value of 0.77, it can be stated that with 100 students as test subjects for mathematical communication instruments, factor analysis is sufficient. Apart from that, the chi-

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.71234

square test is 621.8964, and the p-value is 0.0001, so it can be analyzed using factor analysis. The results of factor analysis can be seen in the Figure 2.

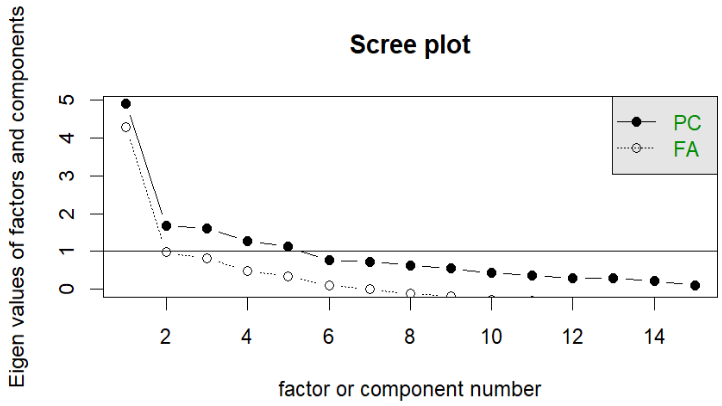

Figure 2. Factor Analysis Results

From the scree plot for factor analysis, 3 factors are obtained because the line starts to slope after three factors. Of these factors, there is 1 factor that makes the most dominant contribution, which is quite large for the given variance component. It shows that Mathematical communication instruments measure at least three factors, with 1 factor being the dominant factor. The loading factor values of the three factors are presented in Figure 3.

	PA1	PA3	PA2	h2	u2	com
1	0.76	0.33	0.03	0.692	0.31	1.4
2	0.79	0.34	-0.07	0.748	0.25	1.4
3	0.66	0.23	0.00	0.482	0.52	1.2
4	0.11	0.11	0.78	0.630	0.37	1.1
5	0.10	0.17	0.71	0.540	0.46	1.2
6	0.74	0.11	0.33	0.664	0.34	1.4
7	0.64	0.00	0.29	0.500	0.50	1.4
8	0.28	0.65	0.05	0.281	0.72	1.7
9	-0.02	0.58	0.08	0.084	0.92	1.2
10	0.12	0.59	-0.01	0.099	0.90	1.3
11	0.13	0.58	0.22	0.402	0.60	1.4
12	0.11	0.64	0.27	0.494	0.51	1.4
13	0.15	0.65	-0.06	0.443	0.56	1.1
14	0.34	0.60	0.02	0.368	0.63	1.8
15	0.29	0.58	0.10	0.322	0.68	1.8

Figure 3. Loading Factor Value

This study will consider items with factor loadings > 0.5 in absolute value. Factor 1 consists of questionnaire items 1, 2, 3, 6, and 7. Factor 2 consists of questionnaire items number 4 and 5. Factor 3 consists of questionnaire items 8, 9, 10, 11,12, 13, 14, and 15. So, based on empirical analysis, it is proven that all instrument items are valid for measuring students' mathematical communication skills. Validity is how each item accurately measures the construct to be assessed. Proving Validity can be demonstrated with precision and consistency with the measured results (Hendryadi, 2017).

Reliability of Mathematical Communication Ability Instruments

The reliability of the mathematical communication ability instrument was assessed on all questions because all questions were valid. The results of the reliability estimation of the mathematical communication ability instrument on multiple choice questions with the Cronbach's Alpha test using R Studio are 0.930 > 0.6, so overall, the mathematical

10.21831/pep.v29i1.71234

communication ability instrument on multiple choice questions is reliable. Likewise, to estimate the reliability of the mathematical communication ability instrument on essay questions using the Cronbach's Alpha test using R Studio, namely 0.801 > 0.6, so that overall, the mathematical communication ability instrument on description questions is reliable. Table 4 presents reliability estimates using Cronbach's Alpha test using R Studio.

Table 4. Instrument Reliability

Question Form	*Alpha Cronbach*	Criteria
Multiple choice	0.930	Reliabel
Essay	0.801	Reliabel

From the data processing results, 15 mathematical communication ability test questions consisted of 12 multiple choice questions and three essays. Based on Cronbach's Alpha coefficient, The reliability estimate for the mathematical communication ability instrument in multiple-choice questions is reliable overall. As with the description, overall, it is also reliable. Reliability estimation refers to the measuring instrument's stability and consistency over time (Heale & Twycross, 2015; Sürücü & Maslakçi, 2020). Thus, it can be concluded that overall, the mathematical communication ability instrument is reliable.

Differential Power of Items on Mathematical Communication Skills

The results of different power calculations regarding mathematical communication skills, using Anbuso based on Microsoft Excel are presented in Table 5.

Table 5. Different Power of Question Items

Category	Question Number		Percentage
Good	1,2,3,4,5,6,7,8, 11,12,13,14,15		87%
Enough	9,10		13%
Bad	-		0%
		Amount	100%

Calculating the differentiating power of the items from 12 multiple-choice questions and three description questions, we obtained 13 items in the good category consisting of 10 multiple-choice items in the good category and three in the good category description. Apart from that, there are also two questions, namely multiple choice questions, in the good category. Questions number 9 and 10 are questions in the relatively good category. There are two out of 4 questions with indicators connecting diagrams, graphs, and tables into mathematical ideas. My mathematical communication skills are still quite good. Overall, the differentiating power of questions with indicators connecting diagrams, graphs, and tables into mathematical ideas on mathematical communication skills is still quite good. However, the ability to differentiate questions for indicators using formulas correctly and explaining the reasons for solving problems in their own language has fulfilled the good category.

The discriminating power of questions can be defined as the ability of questions to distinguish students who are in the more intelligent group from students who are in the less intelligent group (Solichin, 2017). From processing 15 questions on mathematical communication skills consisting of 3 indicators, data was obtained showing no questions with poor discriminating power. However, 13 questions (87%) have good discriminating power. Meanwhile, two questions (13%) with reasonably good differences were questions with

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

indicators connecting diagrams, graphs and tables to mathematical ideas. Because there are no questions with poor differential power, the mathematical communication skills questions are a good type of question. Questions that meet the good or very good criteria can categorize students' communication skills between the best and the worst. On the other hand, questions that meet the minimum criteria cannot categorize students between the best and the worst. This is in line with Warju et al. (2020) said that if the criteria for preparing questions are not met, then the questions will not be able to classify students into upper and lower groups.

Level of Difficulty Items on Mathematical Communication Skills

The results of calculating the level of difficulty regarding mathematical communication skills, using Anbuso based on Microsoft Excel, are presented in Table 6.

Table 6. Level of Difficulty Items on Mathematical Communication Skills

Category	Question Number	Percentage
Easy	-	0%
Currently	1,2,3,4,5,6,7,8, 11,12,13,14,15	93%
Difficult	10	7%
	Amount	100%

Calculating the level of difficulty of the items from 12 multiple-choice questions and three description questions, we obtained 14 questions in the medium category, consisting of 11 multiple-choice items in the medium category and three items in the medium category. Apart from that, there is also 1 question, namely a multiple-choice question in the difficult category. Question number 10 is an item in the easy category. There is 1 out of 4 questions with the indicator of connecting diagrams, graphs, and tables into mathematical ideas. Mathematical communication skills are still in the difficult category. So overall, the level of difficulty of questions with indicators connecting diagrams, graphs, and tables to mathematical ideas on mathematical communication skills is still in the medium category. The level of difficulty of the questions for the indicators of using formulas correctly and explaining the reasons for solving the problem in one's own language meets the medium category.

From the data processing results of 15 questions on mathematical communication skills, no questions were found in the easy category. However, there are 14 questions (93%), with a medium level of difficulty. Meanwhile, 1 question (7%) with a difficult level of difficulty is a question with indicators connecting diagrams, graphs, tables to mathematical ideas. If an item has a moderate or sufficient level of difficulty, then the item is considered good (Suryadevara & Bano, 2018). Because they are easy to answer, easy questions are often underestimated by test takers. On the other hand, test takers are encouraged to tackle challenging problems. On the other hand, difficult questions make test takers discouraged from solving them (Kumalasari, 2016). The results of the analysis of the difficulty level of the mathematical communication ability instrument show that question no. The 10 that fall into the difficult category should be changed or discarded. Therefore, every question item, whether difficult or easy, must be revised or discarded.

Ineffective Alternative Answers to Multiple Choice Questions for Mathematical Communication Ability

All questions were answered by a minimum of 5% of test takers, so there were no questions for which alternative answers were ineffective. So it can be concluded that the

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.71234

alternative answers to the three indicators of mathematical communication skills meet the effective criteria.

There are two possible answers to each question: one is the correct answer (answer key), and the other is the wrong answer. Incorrect answers are known as alternative answers or distractors. Alternative answers work well if they are chosen by at least 5% of test takers. From the data processing results, there were no alternative answers that were ineffective for the 12 multiple choice questions on mathematical communication skills. In multiple choice questions, distractors use the alternative answer function to make participants feel unsure or confused about choosing the right answer. An unselected distractor indicates that the problem needs to be fixed or changed. The level of difficulty of the items is influenced by this distractor function. This is in accordance with research by Iskandar and Rizal (2018) that when there are one or two ineffective distractions, the level of difficulty decreases because it is more likely that test takers will choose the right answer.

Final Conclusions on Mathematical Communication Ability Questions

Based on predetermined criteria. Table 7 summarizes the final conclusions on Microsoft Excel-based Anbuso regarding mathematical communication ability questions.

Table 7. Final Conclusion on Mathematical Communication Ability Questions

Category	Question Number		Percentage
Good	1,2,3,4,5,6, 7,8,9, 11,12,13, 14,15		93%
Enough	10		7%
Bad	-		0%
		Amount	100%

The conclusion about the mathematical communication skills questions is that there is 1 question, namely question number 10, which is quite good. Because question number 10 is an indicator of connecting diagrams, graphs, and tables to mathematical ideas on mathematical communication skills, it can be concluded that there are still questions with indicators of connecting diagrams, graphs, and tables to mathematical ideas in the fairly good category. However, the conclusion for the questions with indicators using formulas correctly and explaining the reasons for solving the problem in the form of one's own language meets the good criteria.

From the results of data processing from 15 questions on mathematical communication skills, the conclusion was that 14 questions (93%) were included in the good category. However, there is 1 question (7%) with an indicator that connects diagrams, graphs, and tables to mathematical ideas, which are in the quite good category, so it is best to revise this question.

CONCLUSION

Based on the findings of the item analysis, it shows that the mathematical communication ability questions show that 15 questions meet the valid and reliable requirements. The differentiating power of the questions reached the good category at 87% (13) and the quite good category at 13% (2). However, the bad category is 0%. The difficulty level of the questions reached the difficult category of 7% (1 question item), and the difficulty level of the questions reached the medium category of 93% (14 questions). However, 0% or no questions reach the easy level of difficulty. Each alternative answer functions well because it reaches a minimum of 5% selected. The final conclusion from the 16-item mathematical communication ability instrument was that 93% were accepted, and 7% were changed. 7% (1 question) are questions with indicators that connect diagrams, graphs, and tables to mathematical ideas. The limitation

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

of this research is that it only creates questions on one topic, namely statistics. So, further research needs to be expanded to other mathematical topics.

Conflict of interests

There are no known conflicts of interest associated with this publication.

REFERENCES

Alamsyah, N. (2015). Pengembangan Instrumen Komunikasi Matematis Untuk Siswa SMP. *Research and Development Journal Of Education*, *2*(1), 29–40.

Alfiatunnisa, E., Zulfah Khairunnisa, H., Hayati, S., & Listya Maulida, V. (2022). Uji Validitas Dan Reliabilitas Terhadap Kemandirian Siswa Sekolah Dasar Kelas 1. *JURNAL HURRIAH: Jurnal Evaluasi Pendidikan Dan Penelitian*, *3*(2), 29–36. https://doi.org/10.56806/jh.v3i2.81

Andini, S. F., & Marlina, R. (2021). Analisis Kemampuan Komunikasi Matematis Siswa Smp Dalam Menyelesaikan Soal Pada Materi Himpunan. *Rumus Hitung*, *4*(2), 343–354. https://doi.org/10.22460/jpmi.v4i2.343-354

Azmi, M. P., & Salam, A. (2020). Pengembangan Instrumen Tes Kemampuan Komunikasi Matematis pada Materi Segi Empat. *JURING (Journal for Research in Mathematics Learning)*, *3*(2), 181. https://doi.org/10.24014/juring.v3i2.10029

Babys, U. (2020). Analisis Kemampuan Komunikasi Matematika Siswa Ditinjau dari Gender. *ANARGYA: Jurnal Ilmiah Pendidikan Matematika*, *3*(1), 25–29. https://doi.org/10.24176/anargya.v3i1.4771

Boopathiraj, C., & Chellamani, K. (2013). Analysis of Test Items on Difficulty Level and Discrimination Index in the Test for Research in Education. *International Journal of Social Science & Interdisciplinary Research*, *2*(2), 189–193.

Friatma, A., & Anhar, A. (2019). Analysis of validity, reliability, discrimination, difficulty and distraction effectiveness in learning assessment. *Journal of Physics: Conference Series*, *1387*(1). https://doi.org/10.1088/1742-6596/1387/1/012063

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, *18*(3), 66–67. https://doi.org/10.1136/eb-2015-102129

Hendryadi. (2017). Validitas isi: tahap awal pengembangan kuesioner. *Jurnal Riset Manajemen Dan Bisnis (JRMB) Fakultas Ekonomi UNIAT*, *2*(2), 169–178.

Iskandar, A., & Rizal, M. (2018). Analisis kualitas soal di perguruan tinggi berbasis aplikasi TAP. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *22*(1), 12–23. https://doi.org/10.21831/pep.v22i1.15609

Istiyono, E. (2020). *Pengembangan Instrumen Penilaian dan Analisis Hasil Belajar Fisika dengan Teori Tes Klasik dan Modern (Edisi Kedua). UNY Press* (Vol. 27, Issue 2, pp. 216–225).

Jannah, R., Hidayat, D. N., Husna, N., & Khasbani, I. (2021). An item analysis on multiple-choice questions: a case of a junior high school English try-out test in Indonesia. *Leksika: Jurnal Bahasa, Sastra Dan Pengajarannya*, *15*(1), 9. https://doi.org/10.30595/lks.v15i1.8768

K, I., Jahring, J., & Subawo, M. (2021). Analisis Kemampuan Komunikasi Matematis Ditinjau Dari Gaya Belajar. *Square : Journal of Mathematics and Mathematics Education*, *3*(1), 56–65. https://doi.org/10.21580/square.2021.3.1.7704

10.21831/pep.v29i1.71234

Khadijah, I. N. A., Maya, R., & Setiawan, W. (2018). Kemampuan Komunikasi Matematis Siswa SMP pada Materi Statistika. *Plusminus: Jurnal Pendidikan Matematika*, *1*(3), 505–518. https://doi.org/10.31980/plusminus.v1i3.1446

Kumalasari, E. (2016). Analisis faktor kesulitan terhadap kesalahan penyelesaian soal persamaan linier berdasarkan klasifikasi Taksonomi Bloom (Studi kasus terhadap mahasiswa Teknik Informatika 2015/2016). *Jurnal Penelitian Pendidikan Dan Pengajaran Matematika*, *2*(2), 113–122.

Kuswandi, K., & Astuti, H. P. (2019). Analisis Kemampuan Komunikasi Matematis Siswa melalui Pembelajaran Kooperatif Tipe Jigsaw. *Suska Journal of Mathematics Education*, *5*(1), 47. https://doi.org/10.24014/sjme.v5i1.6695

Magdalena, I., Fauziah, S. N., Faziah, S. N., & Nupus, F. S. (2021). Analisis Validitas, Reliabilitas, Tingkat Kesulitan Dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan. *BINTANG : Jurnal Pendidikan Dan Sains*, *3*(2), 198–214. https://ejournal.stitpn.ac.id/index.php/bintang

NCTM. (2000). Principles and standards for school mathematics: A guide for mathematicians. *Notices of the American Mathematical Society*, *47*(8), 868–876.

Kholis, R. A. N. (2018). Analisis Tingkat Kesulitan (Difficulty Level) Soal Pada Buku Sejarah Kebudayaan Islam. *Jurnal Penelitian Ilmiah INTAJ*, *1*(2), 93–113. https://doi.org/10.35897/intaj.v1i2.111

Nurlaila, S., Sariningsih, R., & Maya, R. (2018). Analisis kemampuan komunikasi matematis siswa SMP terhadap soal-soal bangun ruang sisi datar. *JPMI (Jurnal Pembelajaran Matematika Inovatif)*, *1*(6), 1113–1120.

Qohar, A., & Fazira, S. K. (2022). Student Mathematical Communication in Online Discussion in Introduction to Geometry Course using Edmodo. *Journal of Education Research and Evaluation*, *6*(4), 576–585. https://doi.org/10.23887/jere.v6i4.51241

Revita, R., Kurniati, A., & Andriani, L. (2018). Analisis Instrumen Tes Akhir Kemampuan Komunikasi Matematika Untuk Siswa Smp Pada Materi Fungsi Dan Relasi. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, *2*(2), 8–19. https://doi.org/10.31004/cendekia.v2i2.44

Sa'idah, N., Yulistianti, H. D., & Megawati, E. (2018). Analisis Instrumen Tes Higher Order Thinking Matematika Smp. *Jurnal Pendidikan Matematika*, *13*(1), 41–54. https://doi.org/10.22342/jpm.13.1.6619.41-54

Solichin, M. (2017). Analisis Daya Beda Soal, Taraf Kesukaran, Validitas Butir Tes, Interpretasi Hasil Tes dan Validitas Ramalan dalam Evaluasi Pendidikan. *Dirāsāt: Jurnal Manajemen & Pendidikan Islam*, *2*(2), 192–213. www.depdiknas.go.id/evaluasi-proses-

Supahar, S., & Prasetyo, Z. K. (2015). Pengembangan Instrumen Penilaian Kinerja Kemampuan Inkuiri Peserta Didik Pada Mata Pelajaran Fisika Sma. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *19*(1), 96–108. https://doi.org/10.21831/pep.v19i1.4560

Suryadevara, V. K., & Bano, Z. (2018). Item analysis to identify quality multiple choice questions/items in an assessment in Pharmacology of II MBBS students in Guntur Medical College of Andhra Pradesh, India. *International Journal of Basic and Clinical Pharmacology*, *7*(8), 1517. https://doi.org/10.18203/2319-2003.IJBCP20183004

Sürücü, L., & Maslakçi, A. (2020). Validity and Reliability in Quantitative Research. *Business & Management Studies: An International Journal*, *8*(3), 2694–2726. https://doi.org/10.15295/bmij.v8i3.1540

Wahyuni, V. (2022). Validity and Reliability of Mathematical Communication Ability Test Instruments Relation and Function Materials. *Sustainable Jurnal Kajian Mutu Pendidikan*, *5*(1), 89–99. https://doi.org/10.32923/kjmp.v5i1.2232

Warju, W., Ariyanto, S. R., Soeryanto, S., & Trisna, R. A. (2020). Analisis Kualitas Butir Soal Tipe Hots Pada Kompetensi Sistem Rem Di Sekolah Menengah Kejuruan. *Jurnal Pendidikan Teknologi Dan Kejuruan*, *17*(1), 95. https://doi.org/10.23887/jptk-undiksha.v17i1.22914

Wijayanto, A. D., Fajriah, S. N., & Anita, I. W. (2023). Analisis Kemampuan Komunikasi Matematis Siswa SMP pada Materi Aljabar. *Radian Journal: Research and Review in Mathematics Education*, *1*(3), 127–134. https://doi.org/10.35706/rjrrme.v1i3.7153