DEVELOPMENT OF A TEST INSTRUMENT TO MEASURE MATHEMATICAL COMMUNICATION SKILLS ON THE TOPICS OF THREE-VARIABLE LINEAR EQUATIONS AND TWO-VARIABLE LINEAR INEQUALITIES

Aprillia Anggita Simanullang[1], *Kartini[2], Titi Solfitri[3]
[1,2,3] Universitas Riau, Indonesia
kartini@lecturer.unri.ac.id

ABSTRACT Mathematical communication skills are essential in helping students express, interpret, and solve mathematical problems. However, these skills are often underdeveloped due to a lack of targeted test instruments aligned with relevant mathematical content. This study aims to develop a valid and reliable test instrument to measure mathematical communication skills in the context of three-variable linear equations and two-variable linear inequalities. Using a development research approach with preliminary and formative evaluation stages adapted from Tessmer's model, the instrument was systematically designed, validated, and tested. Data were collected through expert reviews, student responses, and field testing involving 30 participants. Internal validation results indicated a very high validity score (92.80%), while field testing showed that 12 of the 13 developed items were externally valid. The reliability coefficient was 0.80, categorized as very high. The final test set included one easy item, nine medium items, and two difficult items, with discriminating power ranging from sufficient to very good. These findings demonstrate that the developed test instrument is both valid and reliable, and it effectively assesses mathematical communication skills with an appropriate balance of difficulty and item discrimination.

Keywords: test development, mathematical communication, linear equations, linear inequalities, item validation

ABSTRAK Kemampuan komunikasi matematis merupakan keterampilan penting yang membantu peserta didik dalam mengungkapkan, menafsirkan, dan menyelesaikan permasalahan matematika. Namun, kemampuan ini masih belum berkembang secara optimal akibat minimnya instrumen tes yang secara spesifik dirancang untuk mengukur aspek komunikasi dalam konteks materi matematika tertentu. Penelitian ini bertujuan untuk mengembangkan instrumen tes yang valid dan reliabel guna mengukur kemampuan komunikasi matematis pada materi sistem persamaan linear tiga variabel dan pertidaksamaan linear dua variabel. Penelitian menggunakan pendekatan penelitian

 https://doi.org/10.30606/absis.v8i1.2745

pengembangan dengan tahapan preliminary dan formative evaluation yang diadaptasi dari model Tessmer. Instrumen dikembangkan secara sistematis melalui validasi ahli, respon peserta didik, dan uji coba lapangan terhadap 30 responden. Hasil validasi internal menunjukkan tingkat kevalidan yang sangat tinggi (92,80%), dan pada tahap uji lapangan diperoleh 12 dari 13 butir soal valid secara eksternal. Nilai reliabilitas mencapai 0,80 dengan kategori sangat tinggi. Instrumen akhir terdiri atas satu soal mudah, sembilan soal sedang, dan dua soal sukar, dengan daya pembeda berkisar dari cukup hingga sangat baik. Hasil ini menunjukkan bahwa instrumen yang dikembangkan telah memenuhi kriteria validitas dan reliabilitas, serta mampu mengukur kemampuan komunikasi matematis secara efektif dengan tingkat kesukaran dan daya pembeda yang sesuai.

Kata-kata kunci: pengembangan tes, komunikasi matematis, persamaan linear, pertidaksamaan linear, validasi butir soal.

INTRODUCTION

Mathematical communication skills are essential competencies that students must possess in learning mathematics. These skills involve conveying mathematical ideas both orally and in writing, using symbols or mathematical language to represent ideas in the form of diagrams, graphs, and tables, as well as explaining and constructing arguments related to mathematical problems (Asuro & Fitri, 2020). In the learning process, such skills help students understand the material and solve mathematical problems (Hadijah, Hasratuddin, & Napitupulu, 2016). Good mathematical communication also contributes to constructing mathematical models needed to solve problems across scientific disciplines and in everyday life (Hendriana & Soemarmo, 2016).

According to Asikin (in Rizqi, 2016), mathematical communication can sharpen students' thinking in understanding the connections between mathematical topics. Through communication, students are able to formulate and reinforce their mathematical reasoning. It also serves as a tool to explore mathematical ideas, assess and reflect on students' understanding, organize and consolidate mathematical thinking, and construct mathematical knowledge. Furthermore, it supports problem solving, enhances reasoning and self-confidence, and develops social skills (Noviyana et al., 2019). The presence of such skills is crucial for encouraging active and enjoyable learning and for eliminating the perception that mathematics is a difficult and intimidating subject (Choridah, 2013).

The 2013 Curriculum emphasizes the importance of mathematical communication in students' competencies, such as expressing ideas through symbols, tables, diagrams, or other media to clarify situations or problems, as well as developing logical, analytical, systematic, critical, and creative thinking abilities (Putri & Santosa, 2015). Likewise, the Independent Curriculum emphasizes the development of independence, critical reasoning, and student creativity (Kemendikbud, 2022). According to the NCTM process standards, mathematical communication is one of the fundamental mathematical abilities, alongside problem-solving, reasoning, connections, and representation. Sari (2020) stated that mathematical communication involves organizing mathematical thinking, logically and clearly communicating ideas, evaluating others' reasoning, and using mathematical language accurately. These activities foster high-order thinking skills such as logical, systematic, critical, creative, and productive thinking.

However, several studies have shown that students' mathematical communication skills are still relatively low. Ermita (2021) found that Grade XI students at SMA N Kampar Timur had difficulty solving two-variable linear inequality problems, with a 50% error rate on indicators related to drawing graphs. Similarly, Usman et al. (2021) reported that students struggled to explain mathematical models and understand inequality graphs. Lelboy et al. (2021) also found that students were unable to construct mathematical models and did not understand the steps to solve problems, with an error rate of up to 69%.

This decline is suspected to be due to the lack of mathematical communication exercises provided by teachers. Muir (in Dwi et al., 2018) noted that students often fail to create mathematical models from contextual problems. Hikmawati et al. (2019) emphasized the need for regular practice to develop mathematical communication skills. However, Ansari (in Musfiya et al., 2022) found that teachers tended to explain problems rather than assign exercises that promote communication. This is supported by Sudrajat (2022), who noted that the exercises given are usually routine and do not encourage mathematical communication.

Interviews with mathematics teachers from three high schools in Pekanbaru revealed that teachers have not yet provided items that explicitly assess indicators of mathematical communication. They indicated that the available items were very limited and often reused from textbooks or instructional materials. Furthermore, they were uncertain whether the items they developed qualified as mathematical communication problems.

Some previous studies have attempted to develop test instruments for assessing mathematical communication skills, but they have not addressed all aspects of validity and item quality. For example, Azmi (2020) developed four items on quadrilateral material that were valid and reliable, but did not undergo internal validation testing. Wahyuni (2022) developed 14 items on relations and functions that were internally and externally valid, but were not tested for difficulty level and discriminating power. Therefore, the development of instruments that fulfill all key aspects, such as validity, reliability, difficulty level, and discriminating power, is urgently needed.

The instrument developed in this study is expected to serve as a reference for teachers to effectively train and assess students' mathematical communication skills. The development focuses on the topics of three-variable linear equations and two-variable linear inequalities, with open-ended items designed to reveal various relevant indicators of mathematical communication.

METHODS

 https://doi.org/10.30606/absis.v8i1.2745

This study employed development research using the *development studies* approach adapted from Tessmer (in Dewi & Syofiana, 2020). The development process consisted of two main stages: preliminary and formative evaluation. The preliminary stage involved analysis and initial design, while the formative evaluation included five phases: self-evaluation, expert review, one-to-one, small group, and field test. The overall product development flow is illustrated in Figure 1.

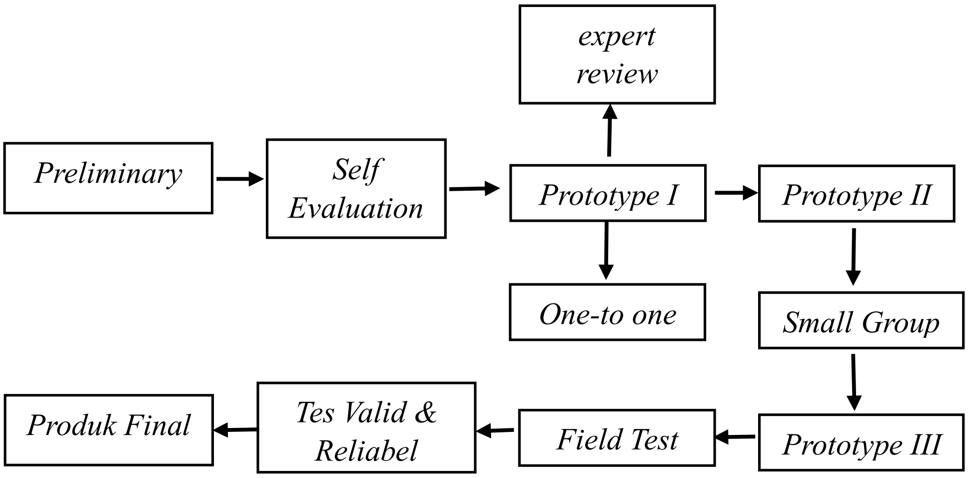

Figure 1. Flowchart of the Development Stages of the Mathematical Communication Skills Test Instrument

In the preliminary stage, a needs analysis, curriculum analysis, and learner analysis were conducted. The needs analysis was based on literature review, interviews with mathematics teachers of Grade X in three high schools in Pekanbaru, and a document review of assessment instruments. Learner analysis considered cognitive characteristics of high school students aligned with the independent curriculum, while curriculum analysis reviewed the learning objectives for the topics of three-variable linear equations and two-variable linear inequalities.

The results of the analysis informed the design stage, which involved developing the initial product, including the test blueprint, open-ended items, scoring rubrics, validation sheets, and student response questionnaires. This initial product was reviewed through self-evaluation to ensure content completeness and appropriateness, resulting in Prototype I.

Prototype I was validated by three mathematics education experts during the expert review phase and was also tested individually on three students of varying ability levels (high, medium, low) during the one-to-one phase. Feedback from both experts and students was used to revise the instrument into Prototype II.

In the small group phase, six students representing different ability levels completed the instrument and provided feedback on item clarity and readability. Revisions based on this feedback produced Prototype III, which was then administered in the field test phase to 30 students with heterogeneous abilities.

https://doi.org/10.30606/absis.v8i1.2745

Data from the field test were analyzed to determine item validity (using Pearson correlation), reliability (using Cronbach's Alpha), difficulty level, and discriminating power. Internal validity was established through expert judgment using validation sheets, covering content, construction, and language aspects, with a Likert scale (1–5). An item was considered internally valid if it scored at least in the categories of moderately valid (40–60%), valid (60–80%), or highly valid (80–100%).

External validity was assessed through student responses and statistical analysis, including correlation and t-tests. Items that met the criteria for external validity were further analyzed for reliability, which was categorized as high (0.60–0.80) or very high (0.80–1.00). Difficulty levels were classified as difficult, moderate, or easy. Discriminating power was also calculated and categorized as sufficient, good, or very good, based on standard criteria (Hamzah, 2014). Thus, each final item was considered valid, reliable, of appropriate difficulty, and capable of effectively distinguishing students' levels of ability.

FINDING AND DISCUSSION

This section presents the research findings based on the instrument development stages, starting from the preliminary stage to the field test, along with discussions on the validity, reliability, difficulty level, and discriminating power of the test items.

Preliminary Stage

The needs analysis showed that mathematical communication skills are essential to be developed in mathematics learning. Three key findings were obtained from this analysis: (1) The questions used by teachers generally had a low level of difficulty and did not adequately train students' mathematical communication skills, (2) The availability of diverse questions was very limited and tended to be sourced from textbooks or similar teaching materials, (3) Teachers were not clearly aware of the indicators of mathematical communication in the questions given to students.

Curriculum analysis showed that the content of three-variable linear equations and two-variable linear inequalities in Phase E learning outcomes required students to solve contextual problems using mathematical concepts and skills. The student analysis indicated that Grade X students aged 16–17 were cognitively able to think abstractly and symbolize information, making them suitable for solving contextual mathematical communication questions.

Design Stage

Based on the analysis, 13 open-ended items were designed to include three mathematical communication indicators: (1) expressing situations into mathematical models and solving them (mathematical expression), (2) representing situations through graphs (drawing), and (3) explaining mathematical ideas or relationships in writing (written text). The distribution of indicators across the 13 questions includes: items 1–11a reflect mathematical expression; items 6b, 7b, 8b, 9b, 10b, and 11b

 https://doi.org/10.30606/absis.v8i1.2745

reflect drawing abilities; and items 3b, 5b, 10c, 11c, 12, and 13 reflect written explanation.

An example of one of the developed items is:

Item Number 1: Laura, Tiara, and Ori are visiting a bazaar. They want to determine the unit prices of mini martabak, churros, and coconut jelly based on three different packages:

Package 1: Rp50,000, contains 2 mini martabak, 1 churros, 1 coconut jelly

Package 2: Rp56,000, contains 1 mini martabak, 2 churros, 1 coconut jelly

Package 3: Rp58,000, contains 1 mini martabak, 1 churros, 2 coconut jelly

Determine the unit price of each item.

This item measures the ability to construct a system of three-variable linear equations from contextual information and solve it, representing the mathematical expression indicator.

Self-Evaluation and Expert Review

In this stage, the 13 developed items were reviewed by three validators, who were mathematics education lecturers. The validation results showed an average score of 92.80%, categorized as very valid. The validation details are shown in Table 1.

Table 1. Validation Results of the Mathematical Communication Skills Test Instrument

Aspect	Content	Construction	Language
Score	63.57	62.5	55.08
Average	4.89	4.80	4.23
Percentage	97.80	96.00	84.73
Category	Very Valid	Very Valid	Very Valid

Table 7 shows that the validation percentages for content, construction, and language aspects were all in the very valid category, with scores of 97.80%, 96.00%, and 84.73% respectively. This indicates that the developed items were highly relevant to the mathematical communication indicators, structurally well-organized, and linguistically appropriate for the students' level. With a total validation score of 92.80%, the instrument is considered suitable for use after minor revisions.

The validators also provided technical improvement suggestions, which included refining wording and typographical accuracy, adjusting working time, avoiding repetition of indicators across different items, and enhancing the clarity of language and terminology in the instructions and scoring guides. These suggestions are summarized in Table 2.

Table 2. Summary of Suggestions from Validators

Validator	Component	Summary of Suggestions

Validator 1	Test Items	Improve editorial accuracy and formatting consistency in the questions.
Validator 2	Test Items	Review completion time and avoid using the same indicators repeatedly.
Validator 3	General	Enhance language use, punctuation, and consistency in technical terms and instructions.

Table 2 indicates that most of the suggestions were technical and not related to the substance of the items. This confirms that the instrument was conceptually sound, with revisions mainly required for clarity, language, and presentation. These refinements made the instrument easier to understand and more effective.

One-to-One Stage

The prototype was tested on three students of varying abilities. They noted that some item narratives were quite lengthy but still understandable. They suggested presenting questions using tables or visuals to enhance comprehension. Time constraints were also noted, especially for items with complex contexts.

Small Group Stage

The instrument was tested on six students with diverse ability levels. Results showed that the items were well understood and students could plan their solutions effectively. The average student response score was 82.85%, falling into the very good category. Revisions were made to improve item wording and clarity of instructions.

Field Test Stage

The field test involved 30 students to assess item validity, reliability, difficulty level, and discriminating power. Results indicated that 12 of the 13 items were externally valid (t-count > t-table), while one item was invalid due to students' inability to model the problem properly. The results are detailed in Table 3.

Table 3. Item Validity Test Results

Item No.	t-count	t-table	Category
1	4.01	2.05	Valid
2	4.63	2.05	Valid
3	3.77	2.05	Valid
4	5.70	2.05	Valid
5	2.63	2.05	Valid
6	2.99	2.05	Valid
7	1.20	2.05	Not Valid
8	3.62	2.05	Valid
9	3.62	2.05	Valid

https://doi.org/10.30606/absis.v8i1.2745

Item No.	t-count	t-table	Category
10	3.31	2.05	Valid
11	4.73	2.05	Valid
12	4.68	2.05	Valid

According to Table 9, 12 out of 13 items were empirically valid with t-count values exceeding the t-table. Only item number 7 was considered invalid, likely due to its complexity or ambiguity, which hindered students' ability to form a correct mathematical model. Therefore, the item was excluded from the final instrument.

The 12 valid items were then analyzed for reliability using Cronbach's Alpha, resulting in a coefficient of 0.80, which falls under the very high category. The difficulty level analysis showed 1 easy item, 9 moderate items, and 2 difficult items. Discriminating power analysis revealed 4 items with sufficient discrimination, 7 with good discrimination, and 1 with very good discrimination. These results indicate that the developed items are of good quality and effectively distinguish between students of different ability levels.

Another example representing the "drawing" indicator is:

Item Number 6b: A shop worker offers gift-wrapping services. Gift A requires 1 sheet of paper and 1 meter of ribbon. Gift B requires 1 sheet of paper and 2 meters of ribbon. If there are 12 sheets of paper and 16 meters of ribbon available, create a mathematical model and graph the solution region.

This item measures the ability to construct a system of inequalities and represent the feasible region graphically. Based on the analysis, it has moderate difficulty and sufficient discriminating power.

Overall, the final instrument consists of 12 items that are proven to be valid and reliable, meeting the criteria for item difficulty and discriminating power. The superiority of this instrument lies in its empirical quality, having passed internal and external validation, reliability analysis, readability checks, and item characteristic analysis. This makes the instrument a precise and consistent tool for assessing students' mathematical communication skills, and it can be effectively used by teachers in classroom assessment.

This advantage distinguishes the present study from prior research by Azmi (2020) and Wahyuni (2022), which did not demonstrate external validity, reliability, or overall item quality due to limited time that prevented the implementation of the small group testing phase.

CONCLUSIONS AND RECOMMENDATIONS

Based on the results of this study, it can be concluded that the mathematical communication skills test instrument developed through the preliminary and formative evaluation stages has met the criteria for validity, reliability, item difficulty, and discriminating power. Of the 13 items developed, 12 were empirically validated and found to be reliable, with a reliability coefficient of 0.80. These items also demonstrated a proportional distribution of difficulty levels and adequate discriminating power. The instrument is considered suitable for use as an assessment tool in mathematics learning, as it accurately, consistently, and meaningfully measures students' mathematical communication skills. Therefore, teachers can utilize this instrument for both formative and diagnostic assessment purposes, particularly in the topics of three-variable linear equations and two-variable linear inequalities. Future researchers are encouraged to develop similar instruments for other mathematical topics and test their applicability across different educational levels and contexts to enhance generalizability. Additionally, the findings of this development can serve as a reference for curriculum developers and policymakers in designing assessments that align with the goals of the independent curriculum and support the strengthening of students' mathematical communication competencies.

REFERENCES

Asuro, N., & Fitri, I. (2020). Analisis kemampuan komunikasi matematis ditinjau dari self concept siswa SMA/MA Nur. *Suska Journal of Mathematics Education, 6*(1), 33–46.

Azmi, M. P., Salam, A., Kasim, S., & Studi Pendidikan Matematika UIN Sultan. (2020). Pengembangan instrumen tes matematis pada materi segi empat kemampuan komunikasi. *Jurnal Pendidikan*, *3*(3), 181–192.

Choridah, D. T. (2013). Peran pembelajaran berbasis masalah untuk meningkatkan kemampuan komunikasi dan berpikir kreatif serta disposisi matematis siswa SMA. *Infinity Journal, 2*(2), 194. https://doi.org/10.22460/infinity.v2i2.35

Ermita, E. (2021). Peningkatan pemahaman program linear melalui pembelajaran diagnosa dan remedial pada siswa kelas XI IPA 6 SMAN 1 Kampar Timur tahun pelajaran 2020. *Jurnal Pendidikan Tambusai, 5*, 8364–8374. https://www.jptam.org/index.php/jptam/article/view/2339

Darma, B. (2021). *Statistika penelitian menggunakan SPSS: Uji validasi, uji reliabilitas, regresi linier sederhana, regresi linier berganda, uji t, uji f, dan uji r*. Jakarta: Guepedia.

Desmita. (2009). *Psikologi perkembangan peserta didik*. Bandung: Remaja Rosdakarya.

Dewi, A. P., & Syofiana, M. (2020). Pengembangan soal kemampuan penalaran matematis siswa kelas VII SMP. *Gammath: Jurnal Ilmiah Program Studi Pendidikan Matematika, 5*(2), 109–116. http://jurnal.unmuhjember.ac.id/index.php/JPM/article/view/5359

 https://doi.org/10.30606/absis.v8i1.2745

Dwi, M., Edy, B., & Irawati, S. (2018). Kemampuan pemecahan masalah kontekstual siswa SMA pada materi barisan dan deret. *Jurnal Pendidikan, 3*(5), 671–678.

Hadijah, S., Hasratuddin, & Napitupulu, E. (2016). Pengaruh pembelajaran kooperatif tipe Jigsaw terhadap kemampuan pemahaman konsep dan komunikasi matematik siswa SMP Negeri 4 Percut Sei Tuan. *Jurnal Tabularasa*, 285–299.

Hamzah, A. (2014). *Evaluasi pembelajaran matematika*. Jakarta: Rajawali Pers.

Hendriana, H., & Soemarmo, U. (2016). *Penilaian pembelajaran matematika*. Bandung: Refika Aditama.

Hikmawati, N. N., Nurcahyono, N. A., & Balkist, P. S. (2019). Kemampuan komunikasi matematis siswa dalam menyelesaikan soal geometri kubus dan balok. *PRISMA, 8*(1), 68. https://doi.org/10.35194/jp.v8i1.648

Kemendikbud RI. (2022). *Keputusan Kepala Badan Standar, Kurikulum dan Asesmen Pendidikan Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi No. 033 tahun 2022 tentang Capaian Pembelajaran pada Pendidikan Anak Usia Dini, Jenjang Pendidikan Dasar, dan Jenjang Pendidikan Menengah*.

Lelboy, N. E., Nahak, S., & Simarmata, J. E. (2021). Analisis kesalahan dalam menyelesaikan soal matematika sistem persamaan linear tiga variabel. *MES: Journal of Mathematics Education and Science, 7*(1), 10–20. https://doi.org/10.30743/mes.v7i1.4347

Musfiya, Nasriadi, A., & Amalia, Y. (2022). Analisis kemampuan komunikasi matematis siswa pada materi SPLDV (Sistem Persamaan Linear Dua Variabel) kelas X di SMAN 11 Banda Aceh. *Jurnal Ilmiah Mahasiswa Pendidikan, 3*(2).

Noviyana, I. N., Dewi, N. R., & Rochmad, R. (2019). Analisis kemampuan komunikasi matematis siswa ditinjau dari self-confidence. *PRISMA, Prosiding Seminar Nasional Matematika, 2*, 704–709.

Nuryadi, & Khuzaini, N. (2016). *Evaluasi hasil dan proses pembelajaran matematika*. Yogyakarta: LeutikaPrio.

Putri, R. I., & Santosa, R. H. (2015). Keefektifan strategi REACT ditinjau dari prestasi belajar, kemampuan penyelesaian masalah, koneksi matematis, self-efficacy. *Jurnal Riset Pendidikan Matematika, 2*(2), 262–272. https://doi.org/10.21831/jrpm.v2i2.7345

Rizqi, A. A. (2016). Kemampuan komunikasi matematis siswa melalui blended learning berbasis pemecahan masalah. *PRISMA, Prosiding Seminar Nasional Matematika, 1*(1), 191–202. https://journal.unnes.ac.id/sju/index.php/prisma/article/view/21457

Sa'adah, S. I., Rasmiwetti, & Linda, R. (2019). Pengembangan soal HOTS dengan Wondershare Quiz Creator sebagai media display pada materi stoikiometri kelas X. *Jurnal Tadris Kimiya, 4*(2), 177–188.

Sari, I. (2020). Pengaruh penerapan model pembelajaran problem posing terhadap kemampuan komunikasi matematis berdasarkan self-efficacy siswa SMP/MTs (Skripsi, Universitas Islam Negeri Sultan Syarif Kasim Riau).

Usman, M. R., Rifaq, A., & Patalo, G. (2021). Analisis kemampuan komunikasi siswa pada materi program linear. *Jurnal Pendidikan dan Sains Matematika, 7*(2).

Wahyuni, V. (2022). Validitas dan reliabilitas instrumen tes kemampuan komunikasi matematis materi relasi dan fungsi. *Sustainable: Jurnal Kajian Mutu Pendidikan, 5*(1), 89–99. https://doi.org/10.32923/kjmp.v5i1.2232