AKSIOMA: Jurnal Program Studi Pendidikan Matematika
Volume 14, No. 3, 2025, 958-970

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

DEVELOPMENT OF AKM-BASED QUESTIONS TO FACILITATE STUDENTS' NUMERACY SKILLS

Bayu Irwandi[1], Nahor Murani Hutapea[2*], Putri Yuanita[3]

[1,2,3] Universitas Riau, Pekanbaru, Indonesia
**Corresponding author. Pascasarjana Pendidikan Matematika, 28293, Pekanbaru, Indonesia.*
E-mail: bayu.irwandi7426@grad.unri.ac.id [1)]
nahorm.hutapea@lecturer.unri.ac.id [2*)]
putri.yuanita@lecturer.unri.ac.id [3*)]

Received 14 December 2024; Received in revised form 13 August 2025; Accepted 18 September 2025

Abstract

This study was motivated by the importance of numeracy as a core competency required to face 21st-century challenges. However, results of the Minimum Competency Assessment (AKM) reveal that more than half of vocational high school (SMK) students in Indonesia fall below the minimum proficiency level in numeracy. This condition occurs due to the limited availability of practice items for students and the limited ability of teachers to develop AKM-based questions. The aim of this study was to develop valid and reliable AKM-oriented numeracy items to support students' numeracy improvement. The development model employed was development studies, consisting of two stages: the preliminary stage (analysis and design) and the formative evaluation stage. The research was conducted at SMK Abdurrab Pekanbaru using test and interview methods for data collection. During the formative evaluation stage, the developed items were tested on 32 students to assess validity, reliability, difficulty level, and discriminating power. The results showed that 18 test items were valid, comprising 2 multiple-choice, 4 true/false, 3 matching, and 9 essay items. Reliability testing using the KR-20 formula for objective items produced a coefficient of 0.847 (very high), while Cronbach's Alpha for essay items was 0.873 (very high). Difficulty analysis indicated 3 easy, 14 moderate, and 1 difficult item. Discrimination analysis showed that 15 items had good discriminating power and 3 had moderate levels. In conclusion, the 18 developed items were considered suitable for measuring and improving students' numeracy skills and can serve as a reference for teachers in designing AKM-based assessments.

***Keywords**: AKM-based question ;number; numeracy ability*

Abstrak

Penelitian ini dilatarbelakangi oleh pentingnya kemampuan numerasi sebagai kompetensi inti dalam menghadapi tantangan abad ke-21. Namun, hasil Asesmen Kompetensi Minimum (AKM) menunjukkan bahwa lebih dari separuh siswa SMK di Indonesia berada di bawah standar kompetensi minimum dalam numerasi. Kondisi ini terjadi karena terbatasnya ketersediaan soal latihan bagi siswa serta rendahnya kemampuan guru dalam merancang soal berbasis AKM. Penelitian ini bertujuan mengembangkan butir soal numerasi berbasis AKM yang valid dan reliabel untuk melatih kemampuan numerasi siswa. Model pengembangan yang digunakan adalah *development studies* dengan dua tahap: tahap awal (analisis dan desain) dan tahap *formative evaluation*. Penelitian dilaksanakan di SMK Abdurrab Pekanbaru dengan teknik pengumpulan data berupa tes dan wawancara. Pada tahap *formative evaluation*, produk diuji coba kepada 32 siswa untuk menilai validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal. Hasil penelitian menghasilkan 18 butir soal yang valid, terdiri dari 2 soal pilihan ganda, 4 soal benar/salah, 3 soal menjodohkan, dan 9 soal uraian. Uji reliabilitas menggunakan rumus KR-20 untuk soal objektif menghasilkan koefisien 0,847 (kategori sangat tinggi), dan reliabilitas soal uraian menggunakan *Cronbach's Alpha* sebesar 0,873 (kategori sangat tinggi). Analisis kesukaran menunjukkan 3 soal mudah, 14 sedang, dan 1 sukar. Uji daya pembeda menunjukkan 15 soal berkategori baik dan 3 soal cukup. Kesimpulannya, 18 soal yang dikembangkan dinyatakan layak digunakan untuk mengukur dan melatih kemampuan numerasi siswa, serta dapat menjadi acuan bagi guru dalam menyusun asesmen numerasi berbasis AKM.

Kata kunci: Soal berbasis AKM; bilangan; kemampuan numerasi.

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

INTRODUCTION

Mathematical literacy skills, often referred to as numeracy literacy skills, are an essential component in supporting work procedures across various fields. Strong numeracy skills shape individuals who think systematically, analytically, and are able to make accurate decisions in solving various life problems (Wardhani & Oktiningrum, 2022). These skills are part of the essential 21st-century competencies that students must possess to remain competitive in the era of the Industrial Revolution 4.0 (Sumarni et al., 2023). Furthermore, Wijaya & Dewayani (2021) state that understanding students' numeracy skills is an important first step for teachers in designing effective learning. By knowing the level of students' numeracy skills, teachers can adapt teaching methods and materials to suit students' needs.

Based on the 2022 Indonesian Education Report, the numeracy proficiency of vocational high school (SMK) students in Riau Province remains unsatisfactory, with more than 50% of students performing below the minimum competency level (Kementerian Pendidikan, 2022). This indicates that the majority of vocational high school students in this province have not yet reached the expected numeracy standards. This is not much different from the numeracy achievement results of vocational high school students at the national level in 2022, which also fell below the minimum competency level. Quoted from the kemendikbud.go.id website, Minister of Education Mr. Nadiem stated that improving numeracy competence indeed requires a process and cannot be achieved through tutoring alone; what schools need to do is enhance teachers' capacity in delivering lessons and designing quality assessments.

Based on interviews with mathematics teachers at SMK Abdurrab Pekanbaru, it was found that students' numeracy skills at the school were categorized as meeting the minimum competency level, with a score of 1.98. This indicates that 66% of students at SMK Abdurrab Pekanbaru possess adequate numeracy skills. Mathematics instruction at the school is generally student-centered, employing instructional models such as Problem-Based Learning (PBL), Project-Based Learning (PJBL), and cooperative learning. The teachers also design both formative and summative assessments. Formative assessments are conducted at the end of each lesson through short essay-type practice questions or quizzes (1–2 questions) and student feedback questionnaires. Summative assessments include end-of-chapter tests, mid-semester tests, and final exams. End-of-chapter assessments typically consist of 5 to 10 essay questions, while mid-semester tests use multiple-choice questions that require students to show their problem-solving steps. Final exams consist of 40 multiple-choice questions requiring only the selection of the correct answer, without showing the steps

Interviews with mathematics teachers at SMKN 2 Pekanbaru and SMK Perbankan Riau revealed that both schools have sought to enhance students' numeracy skills through the implementation of various innovative learning models, such as PBL, discovery learning, PJBL, and cooperative learning. The findings indicate that innovative teaching strategies hold substantial potential in strengthening students' numeracy

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

competencies (Siregar et al., 2024). However, the absence of adequate assessment instruments remains an obstacle to comprehensively evaluating students' numeracy progress.

Anggraini & Setianingsih (2022) highlight the need to measure education quality using the AKM instrument to guide strategies for improving numeracy skills, which impacts learning outcomes and processes. AKM-based questions are designed to foster understanding, critical thinking, and real-world problem-solving (Jafirah et al., 2024). Accordingly, the present study develops AKM-based questions for vocational students, providing teachers with an effective assessment tool and supporting more tailored learning.

Previous studies have shown limitations in the development of AKM-based items. Mulyani et al. (2024) produced only essay-type questions, with some tested merely for validity, while Sektiwulan & Nindiasari (2024) developed multiple-choice and essay items. On the other hand, Berliana and Berliana & Masriyah (2024) created eight items that were valid and reliable, but these did not include numeracy skill indicators..

Previous studies primarily focused on developing AKM-type items, either in multiple-choice or open-ended formats. However, such studies have not extensively explored the development of daily assessment instruments that adopt the characteristics of AKM questions. Therefore, the present study offers a different approach by constructing daily assessment items modeled after AKM, tailored to specific subject matter. The outcomes of this research are expected to serve as practical references for teachers in designing summative assessments that not only evaluate students' conceptual understanding but also enhance their literacy and numeracy skills in alignment with the objectives of the AKM framework.

This study aims to develop AKM-based assessment items that demonstrate strong validity and reliability, as well as appropriate levels of difficulty and discrimination, to accurately measure vocational high school students' numeracy skills. The final products consist of test specifications, item sets, alternative solutions, and scoring rubrics, all aligned with numeracy indicators. The developed instrument is expected to provide teachers with a practical tool for enhancing both the evaluation process and the effectiveness of learning, thereby contributing to the improvement of students' numeracy skills.

METHODS

This study adopts the development studies model, which consists of a preliminary phase and formative evaluation (prototyping), as introduced by Tessmer (1998). The process involves several key stages, including self-evaluation, expert review, one-to-one testing, small group testing, and field testing.

The preliminary stage in this research consists of two main steps: problem analysis and solution design. The first step, problem analysis, aims to investigate the underlying reasons that necessitate the development of the questions. This process includes needs analysis, student characteristic analysis, and curriculum analysis, which were carried out through classroom observations, interviews with teachers, and document analysis. The second step, solution design, focuses on formulating the solution that becomes

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

the core of this study. At this stage, the researcher develops the initial product or prototype to be further refined and prepares the research instruments required for the subsequent phases.

The formative evaluation stage includes five main steps: self-evaluation, expert review, one-on-one testing, small group, and field testing. In the self-evaluation stage, researchers conduct a self-assessment to evaluate the validity of the product being developed. This process results in the first prototype and the research instruments that will be used in the subsequent stages. This self-evaluation is important to ensure that the initial product meets basic standards before involving experts and other participants in the validation process.

At the expert review stage, three validators assessed the first prototype to obtain internal validity scores and provide constructive suggestions and feedback on the product being developed. This assessment includes three main aspects: content, construction, and language, which are evaluated using a specialized validation sheet. The input and feedback provided by the validators play a crucial role in improving and refining the initial prototype. This assessment serves as the foundation for making revisions and improving the quality of the product. Thus, the final product can meet the expected standards and be ready for further testing in the next stage.

At the one-to-one stage, conducted simultaneously with the expert review stage, the first prototype is directly tested with students. The aim is to gather information regarding the clarity, sequence, completeness, and ease of use of the product, as well as its impact. The input and feedback obtained from the expert review and one-to-one testing serve as the basis for improving the initial prototype. This process results in a second prototype that is better and more aligned with needs. Next, in the small group stage, the second prototype is tested to assess the effectiveness of the changes made based on the revisions from the one-to-one stage and to identify other issues that students might encounter. Feedback from students at this stage is crucial and is used to revise the second prototype, resulting in a third prototype. In the field test stage, the third prototype is tested on a broader subject group to obtain external validity, reliability, difficulty level, and discrimination index. The results of this field test are used to refine the final product, ensuring that it is of high quality and ready for use in actual educational contexts.

This research was conducted at SMK Abdurrab Pekanbaru with tenth-grade students from the *Teknologi Laboratorium Medik* and *Teknologi Farmasi* programs. The selection of research subjects followed several stages based on the development model proposed by Tessmer (1998). In the one-to-one phase, three students from a single class were selected using purposive sampling to evaluate initial comprehension of the developed items. In the small group phase, six students from a different class were also purposively selected to identify common misconceptions and assess item feasibility on a small scale. For the field test phase, cluster sampling was employed by selecting an entire class consisting of 32 students. The selected class met the criteria of having studied the relevant material and being available during the testing period. The selection of this school as the research site was based on its high level of numeracy

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

achievement, as reflected in the 2024 Education Report Card.

The research instruments consisted of a validation sheet and a student response questionnaire. The validation sheet, administered during the expert review stage, assessed internal validity based on three aspects: content, construction, and language. The content aspect examined the alignment of items with learning outcomes, objectives, indicators, cognitive levels, contexts, formats, and alternative solutions in accordance with numeracy indicators—namely, the ability to apply basic mathematical symbols, analyze information in various formats, and interpret results for decision-making. The construction aspect evaluated the appropriateness of item prompts, the suitability of images and tables, and conformity to item types (essay, multiple-choice, true/false, matching). The language aspect ensured proper language use and the absence of ambiguity. The product developed in this study was a set of AKM-based test items, which were later analyzed for validity, reliability, difficulty level, and discrimination power.

The data analysis procedures were carried out in several stages. First, the results of expert validation were analyzed using validation sheets covering content, construct, and language aspects. The scores given by validators on a scale of 1–5 were converted into percentages and then interpreted according to internal validity criteria, namely 81%–100% (very valid), 61%–80% (valid), 41%–60% (fairly valid), 21%–40% (less valid), and 0%–20% (invalid), so that an instrument was considered valid if the percentage score was ≥ 61%. Second, the internal validity of test items was analyzed using the product-moment correlation, where an item was considered valid if the calculated r-value was greater than or equal to the r-table value at the 5% significance level. Third, student response questionnaires were analyzed by converting the obtained scores into percentages, which were then interpreted according to the following criteria: 81%–100% (very good), 61%–80% (good), 41%–60% (fair), 21%–40% (poor), and 0%–20% (very poor). Fourth, the reliability of AKM-based test items was tested, with the instrument deemed reliable if the reliability coefficient fell within the high or very high category (≥ 0.60). Fifth, the difficulty index was calculated and categorized as acceptable if it ranged from medium to difficult ($0 < DI \leq 0.70$). Sixth, the discrimination index was analyzed, with an item considered to have good discriminative power if the index was ≥ 0.20

RESULTS AND DISCUSSION

At the preliminary stage, vocational high school students' numeracy skills were found to be relatively low, with over 50% scoring below the minimum competency level in the 2022 national education report. Interviews with mathematics teachers indicated limited understanding of AKM question characteristics, restricted use of question formats to descriptive and multiple-choice, and difficulties in designing contexts aligned with the taught material or linking alternative solutions to numeracy indicators.

Based on these findings, the product developed consisted of a question grid, alternative solutions, assessment guidelines, and 21 AKM-based questions designed to support numeracy skills, along with validation sheets and student response questionnaires as research instruments.

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

The AKM-based questions align with phase E number elements and include 4 scientific, 2 socio-cultural, and 3 personal contexts, with some contexts applied to multiple items. They comprise 11 application, 5 knowledge, and 5 reasoning items in various formats—2 multiple-choice, 6 true/false, 4 matching, and 9 essay questions—with alternative answers and scoring guidelines adapted to numeracy skill indicators. This product can be used as a teacher assessment guide or a student learning resource to enhance numeracy skills.

Formative evaluation comprises five stages: self-evaluation, expert review, one-on-one testing, small group testing, and field testing. Self-evaluation assesses the completeness of the question grid, question cards, alternative solutions, and scoring guidelines, producing the first prototype. Expert review, conducted by three mathematics education lecturers, validates the prototype in terms of content, construct, and language to ensure internal validity and identify areas for improvement. This process ensures that the instrument is accurate, reliable, and meets academic as well as practical standards in mathematics education.

Validation by three validators during the expert review stage confirmed the product's validity in terms of material, construct, and language, as presented in Table 1.

Table 1. Results of Validation Calculation for Material, Construct, and Language Aspects

Validators	Aspects		
	Material	Construct	Language
I	86,64	90,79	87,87
II	86,69	86,02	92,22
III	89,60	94,12	89,17

Based on the validation results in Table 1, it can be concluded that the validity scores given by each validator fall into the very valid category. The scores for the material, construct, and language aspects are within the range of 81–100.

Based on the validator's suggestions, one question indicator needs to be revised with more detailed specifications to align with the competency being measured. In addition, the validator recommended adding images to each question context and replacing the image in question number 7 to match the illustration in its context. The questions revised based on these suggestions are presented in Table 2.

Table 2. Revisions Based on Expert Review and One-to-One Evaluation Results

Before Revision	After Revision
(Indicator for question number 3). Given an illustration describing the characteristics of a bacterium, students are able to calculate the number of bacteria at the end of a specified period.	(Indicator for question number 3). Given a picture and an illustration describing the characteristics of a bacterium, students are able to calculate the number of bacteria at the end of a specified period by applying the properties of exponentiated numbers.
(Question number 1, 2, 4, 8, 12, and 16). ...berilah tanda centrang (√) pada kolom...	(Question number 1, 2, 4, 8, 12, and 16). ...berilah tanda centang (✓) pada kolom...

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

Before Revision	After Revision
(Question 2 context) *Clostridium botulinum* adalah golongan bakteri basilus anaerobik gram positif…terdapat 32 bakteri clostridium…	(Question 2 context) *Clostridium botulinum* adalah golongan bakteri *basilus anaerobik* gram positif…terdapat 32 bakteri *clostridium*…
(Context Image for Question 4). 	(Context Image for Question 4). 
(Question 8 Context Wording) ...sebesar Rp100 ribu... ...minimal Rp100 ribu. ...minimal Rp25 juta... ...menyetorkan Rp. 500.000,- sebagai ... sebesar Rp. 50.000,- lebih... ...menyetor sebesar Rp. 250.000,- pada bulan Januari dan selalu menambah Rp. 10.000,- dari setoran sebelumnya.	(Question 8 Context Wording) ...sebesar Rp100.000,00... ...minimal Rp 100.000.00. ... minimal Rp 25.000.000,00 ....menyetorkan Rp 500.000,00 sebagai ... sebesar Rp 50.000,00 lebih... ...menyetor sebesar Rp 250.000,00 pada bulan Januari dan selalu menambah Rp 10.000,00 dari setoran sebelumnya.

Alongside expert review, Prototype I was tested with three students in a one-to-one trial, yielding suggestions such as adding problem-solving instructions, page numbers, and improving image quality. Analysis of responses revealed correct reasoning but symbol misuse (e.g., Un vs. Sn), indicating the need for clearer symbol explanations. Feedback informed revisions, producing Prototype II for small group testing.

In the small group stage, students could analyze information but struggled to interpret results for predictions and decisions due to difficulties in using numbers, symbols, and mathematical facts. Questionnaire results showed an 82.86% satisfaction rate (very good), with suggestions to make problem contexts more concise and understandable. All student suggestions and comments from the small group stage are presented in Table 3.

Table 3. Suggestions and Comments from Small Group Stage Students

Suggestions and comments	Revised Decision
High-Ability Sutudents The given problem is interesting to solve because it relates to everyday life. The image presented is black and white and the text is not very large. The context of the presented question can enhance knowledge. There are too many question sheets. Very challenging to work on and practice to be more meticulous. Very interesting because all the questions must be read carefully.	Revising the context of the question to make it simpler.
Middle – Ability Students The time to complete the questions is not enough. Questions number 20 and 21 have not been studied yet.	Unifying the context of the question so that

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

Suggestions and comments	Revised Decision
The context of the question and the question itself are on different pages, so it takes time to flip through the pages.	it fits on one page with the question.
Low – Ability Students In class, we never study the term "general form." The problem is too long. The time provided to answer the questions is not enough. Because there is still a lot to learn.	Revising several formulations of the questions.

Table 4. Results of the Validity Test for True/False, Matching, and Multiple Choice Items

Quesstion Number	r_{hitung}	Category	Quesstion Number	r_{hitung}	Category
1a	0,140	Not Valid	10.1	0,438	Valid
1b	0,530	Valid	10.2	0,428	Valid
1c	0,408	Valid	10.3	0,446	Valid
1d	0,399	Valid	12a	0,536	Valid
1e	0,445	Valid	12b	0,443	Valid
2a	0,408	Valid	12c	0,328	Not Valid
2b	0,426	Valid	12d	0,266	Not Valid
2c	0,410	Valid	12e	0,352	Valid
2d	0,423	Valid	14.1	0,660	Valid
4a	0,433	Valid	14.2	0,446	Valid
4b	0,449	Valid	14.3	0,447	Valid
4c	0,504	Valid	16a	0,275	Not Valid
4d	0,445	Valid	16b	0,401	Valid
4e	0,442	Valid	16c	0,358	Valid
5	0,420	Valid	16d	0,390	Valid
6.1	0,419	Valid	16e	0,491	Valid
6.2	0,199	Not Valid	17	0,442	Valid
6.3	0,225	Not Valid	18.1	0,412	Valid
8a	0,304	Not Valid	18.2	0,371	Valid
8b	0,184	Not Valid	18.3	0,417	Valid
8c	0,198	Not Valid			
8d	0,506	Valid			
8e	0,444	Valid			

Based on Table 3, the product was further revised for testing in the field test stage. In the field test stage, AKM-based questions were trialed on 32 students from class X TF 3 at SMK Abdurrab Pekanbaru to measure the validity of the questions, reliability, difficulty level, and discriminating power of each developed question. Based on the analysis of student answers, the scores ranged from 11 to 93 out of a total score of 106. This score is dominated by the first numeracy skill indicator, which is analyzing information presented in various forms. Although not yet fully developed, students have already demonstrated numeracy skills in the first indicator. These results are in line with the findings of Elina et al. (2022), which show that students' numeracy skills in analyzing information in various forms (such as graphs, tables, charts, and diagrams) received the highest

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

percentage compared to the other two indicators.

The highest-scoring items were question 4 (statement c) and question 12 (statement e). Question 4, in a socio-cultural context at the knowing level, used a complex true/false multiple-choice format to assess the ability to identify characteristics of an arithmetic sequence through seating arrangement illustrations. Question 12, in a scientific context at the same cognitive level, used a similar format to assess identification of geometric sequence characteristics through illustrations of amoeba reproduction. These results indicate that students perform better on questions with clear contexts and visual supports

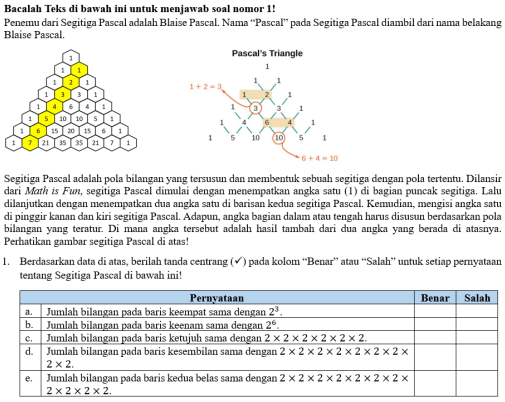
Students demonstrated the ability to identify the characteristics of arithmetic and geometric sequences in AKM-based contexts, consistent with Leonita Bau et al. (2023), who found that students prefer real-life contexts that train their ability to represent numbers and understand concepts. They were also able to analyze information in various forms and solve problems in different contexts (Maulidina & Hartatik, 2019). The validity test of non-essay items using the point-biserial correlation showed that 9 of 44 multiple-choice, true/false, and matching items—namely 1a, 6.2, 6.3, 8a, 8b, 8c, 12c, 12d, and 16a—were invalid.

The validity test of the essay-type questions was calculated using the Pearson product-moment formula, and the result showed that all 9 essay questions were declared valid. The results of the essay question validation are presented in Table 5.

Table 5. Results of the Validity Test for Essay Questions

Quesstion Number	t_{hitung}	Category	Quesstion Number	r_{hitung}	Category
3	8,709	Valid	15	2,768	Valid
7	4,563	Valid	19	11,921	Valid
9	5,458	Valid	20	6,107	Valid
11	8,357	Valid	21	5,872	Valid
13	4,685	Valid			

Questions 1a, 6.2, and 6.3 are examples of invalid items in this study. These items are presented in Figures 1 and 2.



Bacalah Teks di bawah ini untuk menjawab soal nomor 1!

Penemu dari Segitiga Pascal adalah Blaise Pascal. Nama "Pascal" pada Segitiga Pascal diambil dari nama belakang Blaise Pascal.

Pascal's Triangle

Segitiga Pascal adalah pola bilangan yang tersusun dan membentuk sebuah segitiga dengan pola tertentu. Dilansir dari *Math is Fun*, segitiga Pascal dimulai dengan menempatkan angka satu (1) di bagian puncak segitiga. Lalu dilanjutkan dengan menempatkan dua angka satu di barisan kedua segitiga Pascal. Kemudian, mengisi angka satu di pinggir kanan dan kiri segitiga Pascal. Adapun, angka bagian dalam atau tengah harus disusun berdasarkan pola bilangan yang teratur. Di mana angka tersebut adalah hasil tambah dari dua angka yang berada di atasnya. Perhatikan gambar segitiga Pascal di atas!

1. Berdasarkan data di atas, berilah tanda centrang (✓) pada kolom "Benar" atau "Salah" untuk setiap pernyataan tentang Segitiga Pascal di bawah ini!

	Pernyataan	Benar	Salah
a.	Jumlah bilangan pada baris keempat sama dengan 2^3.		
b.	Jumlah bilangan pada baris keenam sama dengan 2^6.		
c.	Jumlah bilangan pada baris ketujuh sama dengan $2 \times 2 \times 2 \times 2 \times 2 \times 2$.		
d.	Jumlah bilangan pada baris kesembilan sama dengan $2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2$.		
e.	Jumlah bilangan pada baris kedua belas sama dengan $2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2$.		

Figure 1. Excerpt from question number 1a

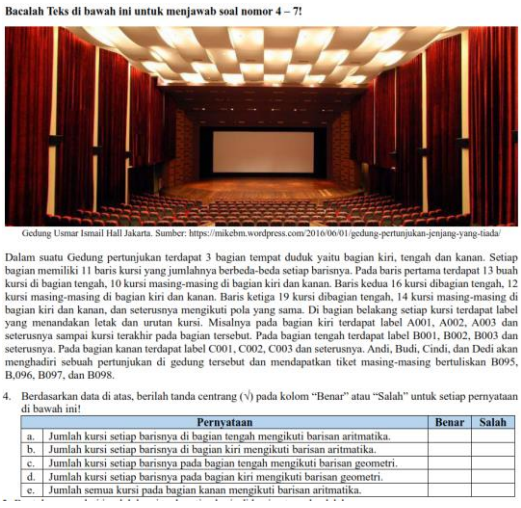
The results for item 1a indicate that the majority of students were able to answer it easily due to its predictable nature. This finding suggests that the level of predictability in a test item may reduce its validity, as the item can no longer effectively distinguish between different levels of student ability (Sudarwan & Retnawati, 2015).

Questions 6.2 and 6.3, in a socio-cultural context at the applying level with a matching format, assessed students' ability to determine the n-th term of an arithmetic sequence. While 6.2 was correctly answered by only 10 students, 6.3 was answered correctly by 27 students. Interviews revealed that high-ability students found the

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

questions manageable, whereas medium- and low-ability students struggled and often guessed, consistent with Wardani et al. (2020) regarding factors causing item invalidity.



Bacalah Teks di bawah ini untuk menjawab soal nomor 4 – 7!

Gedung Usmar Ismail Hall Jakarta. Sumber: https://mikebm.wordpress.com/2016/06/01/gedung-pertunjukan-jenjang-yang-tiada/

Dalam suatu Gedung pertunjukan terdapat 3 bagian tempat duduk yaitu bagian kiri, tengah dan kanan. Setiap bagian memiliki 11 baris kursi yang jumlahnya berbeda-beda setiap barisnya. Pada baris pertama terdapat 13 buah kursi di bagian tengah, 10 kursi masing-masing di bagian kiri dan kanan. Baris kedua 16 kursi dibagian tengah, 12 kursi masing-masing di bagian kiri dan kanan. Baris ketiga 19 kursi dibagian tengah, 14 kursi masing-masing di bagian kiri dan kanan, dan seterusnya mengikuti pola yang sama. Di bagian belakang setiap kursi terdapat label yang menandakan letak dan urutan kursi. Misalnya pada bagian kiri terdapat label A001, A002, A003 dan seterusnya sampai kursi terakhir pada bagian tersebut. Pada bagian tengah terdapat label B001, B002, B003 dan seterusnya. Pada bagian kanan terdapat label C001, C002, C003 dan seterusnya. Andi, Budi, Cindi, dan Dedi akan menghadiri sebuah pertunjukan di gedung tersebut dan mendapatkan tiket masing-masing bertuliskan B095, B,096, B097, dan B098.

4. Berdasarkan data di atas, berilah tanda centrang (√) pada kolom "Benar" atau "Salah" untuk setiap pernyataan di bawah ini!

	Pernyataan	Benar	Salah
a.	Jumlah kursi setiap barisnya di bagian tengah mengikuti barisan aritmatika.		
b.	Jumlah kursi setiap barisnya di bagian kiri mengikuti barisan aritmatika.		
c.	Jumlah kursi setiap barisnya pada bagian tengah mengikuti barisan geometri.		
d.	Jumlah kursi setiap barisnya pada bagian kiri mengikuti barisan geometri.		
e.	Jumlah semua kursi pada bagian kanan mengikuti barisan aritmatika.		

Figure 2 Excerpt from questions number 6.2 and 6.3

The validation results indicated that 18 items were deemed valid (2 multiple-choice, 4 true/false, 3 matching, and 9 essay questions) consisting of 37 statements, while 3 invalid items were excluded from the reliability testing and item analysis. Reliability testing yielded KR-20 = 0.847 for objective items and Cronbach's Alpha = 0.873 for essays, both in the very high category (Riyani et al., 2017). Consistent with Sumarni et al. (2023), these values indicate the AKM-based questions are reliable, producing consistent results across different times and subjects.

The developed AKM-based questions were categorized according to their level of difficulty. The classification results indicated that 9 items fell into the easy category, 24 into the medium category, and 4 into the difficult category. The easy category consisted of 7 true/false questions, 1 matching question, and 1 multiple-choice question. Based on interviews with students of high, medium, and low ability levels, it was revealed that all groups were able to complete the questions successfully. This condition supports the findings of Yuliyanto et al. (2021), which emphasized that students' conceptual readiness plays an important role in reducing their perception of difficulty, thereby making the assessment items appear easier to solve.

In the essay section, 8 questions were identified as having a medium level of difficulty, while 1 question—specifically item number 21—was considered difficult. Student interviews revealed that learners across all ability levels experienced similar challenges with this question. The main source of difficulty was the topic of simple and compound interest, which had been previously taught only through routine tasks, such as calculating the value of savings after a certain period. Furthermore, the question contained several comparative statements, causing students to merely note the given and required information without successfully solving the problem.

Susanto et al. (2015) assumes that, in addition to validity and reliability, the quality of assessment items also depends on a balanced level of difficulty. Thus, the developed questions can be considered high-quality, as they are neither too easy nor too difficult. This aligns with Hidayati & Nisa' (2023), who states that good questions should have a moderate level of difficulty, since overly easy items fail to stimulate students to exert effort in solving problems seriously.

The results of the differential power analysis of the validated AKM-based questions show that there are 25 items categorized as good and 12 items

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

categorized as sufficient. Hanna & Retnawati (2022) states that the discrimination index analysis aims to determine the ability of questions to differentiate between students with high and low abilities. It should also be noted that the results of the difficulty and discrimination index tests may vary depending on students' abilities during the test. If the questions are given to students with a majority of high ability, then the questions tend to have an easy to moderate level of difficulty. Conversely, if the questions are given to students with a majority of low ability, then the questions tend to have a moderate to difficult level of difficulty (Charmila et al., 2016). According Hidayati & Nisa' (2023), questions with poor discrimination power cannot differentiate between students with high and low abilities, so questions with poor discrimination power will be rejected.

In this study, the questions were tested on 32 students from class X TF3 at SMK Abdurrab Pekanbaru. Based on the students' answer scores and referring to the previously described sources, a good question is one that meets the requirements of validity, reliability, and has a good level of difficulty and discrimination. From the results of this research, 18 AKM-based questions were accepted to be used as the final product or question bank to facilitate students' numeracy skills. This product is equipped with steps for developing AKM-based questions, so it can serve as a reference for teachers in creating questions that can measure and train students' numeracy skills. The final product consists of 3 questions on the topic of exponential numbers, 3 questions on the topic of arithmetic sequences, 3 questions on the topic of arithmetic series, 3 questions on the topic of geometric sequences, 4 questions on the topic of geometric series, and 2 questions on the topic of simple and compound interest. This final product includes 4 true/false questions, 3 matching questions, 2 multiple-choice questions, and 9 essay questions.

CONCLUSION AND SUGGESTION

This study concludes that the developed AKM-based numeracy items meet the criteria of high validity, high reliability, and practicality. The final product consists of 18 items representing personal, socio-cultural, and scientific contexts, aligned with numeracy indicators and learning objectives. These results indicate that the developed instrument can support the improvement of vocational students' numeracy skills and serve as a practical reference for teachers in designing AKM-oriented assessments.

Although the developed product has met the criteria of validity and practicality, future development is recommended to incorporate a wider range of difficulty levels, use more engaging visual media such as colored or interactive images, and create a more extensive AKM question bank to accommodate various vocational school majors. Furthermore, subsequent studies could examine the effectiveness of these AKM-based questions in improving students' numeracy skills through experimental or quasi-experimental research designs. Future research could also be extended to other grades or phases, such as Phase F (Grades XI–XII), and explore the integration of digital technology to develop an online version of the AKM-based questions.

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

REFERENCES

Anggraini, K. E., & Setianingsih, R. (2022). Analisis Kemampuan Numerasi Siswa SMA dalam Menyelesaikan Soal Asesmen Kompetensi Minimum (AKM). *Jurnal Ilmiah Pendidikan Matematika*, *11*(3), 837–849. https://doi.org/https://doi.org/10.26740/mathedunesa.v11n3.p837-849

Berliana, A. P., & Masriyah, M. (2024). Pengembangan Soal Model AKM Numerasi Pada Domain Konten Geometri dan Pengukuran Untuk Siswa Kelas VIII SMP. *MATHEdunesa*, *13*(1), 216–233. https://doi.org/10.26740/mathedunesa.v13n1.p216-233

Charmila, N., Zulkardi, Z., & Darmawijoyo, D. (2016). Pengembangan soal matematika model PISA menggunakan Konteks Jambi. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *20*(2), 198–207. https://doi.org/10.21831/pep.v20i2.7444

Elina, E., Maimunah, M., & Roza, Y. (2022). Analysis of the Ability of SMP/MTs Mathematics Teachers in Making AKM Type Questions. *Jurnal Gantang*, *7*(1), 47–57. https://doi.org/10.31629/jg.v7i1.4466

Hanna, W. F., & Retnawati, H. (2022). Analisis Kualitas Butir Soal Matematika Menggunakan Model Rasch Dengan Bantuan Software Quest. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *11*(4), 3695–3704. https://doi.org/10.24127/ajpm.v11i4.5908

Hidayati, K., & Nisa', Z. E. K. (2023). Analisis Butir Soal Penilaian Akhir Semester Matematika. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *12*(3), 3516–3529. https://doi.org/10.24127/ajpm.v12i3.7575

Jafirah, S., Fajriah, N., & Suryaningsih, Y. (2024). Pengembangan Soal Literasi Matematika dengan Konteks Etnomatematika pada Pasar Terapung Untuk Siswa Tingkat SMP/MTs. *EDU-MAT: Jurnal Pendidikan Matematika* , *12*(2), 393–403. https://doi.org/http://dx.doi.org/10.20527/edumat.v12i2.19007

Kementerian Pendidikan (2022). *Rapor Pendidikan Indonesia*. https://data.kemendikdasmen.go.id/publikasi/pendidikan/asesmen-nasional-and-rapor-pendidikan?sub-topik=rapor-pendidikan-indonesia

Leonita Bau, M., Sulistyowati, F., Irfan, M., Ayuningtyas, A. D., & Kusmaningrum, B. (2023). Analisis Kemampuan Numerasi Siswa pada Materi Bilangan Bulat Ditinjau dari Kemandirian Belajar. *SEMANTIK: Prosiding Seminar Nasional Pendidikan Matematika*, 284–292. https://seminar.ustjogja.ac.id/index.php/SEMANTIK/article/view/1896/930

Maulidina, A. P., & Hartatik, S. (2019). Profil Kemampuan Numerasi Siswa Sekolah Dasar Berkemampuan Tinggi Dalam Memecahkan Masalah Matematika. *Jurnal Bidang Pendidikan Dasar (JBPD)*, *3*(2), 61–66. https://doi.org/https://doi.org/10.21067/jbpd.v3i2.3408

Mulyani, S., Lusiana, L., & Jumroh, J. (2024). Pengembangan Soal Asesmen Kompetensi Minimum (AKM) Matematika Pada Materi Peluang Untuk Peserta Didik SMP. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *13*(3), 1054–1063. https://doi.org/10.24127/ajpm.v13i3.8409

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v14i3.11519

Riyani, R., Maizora, S., & Hanifah. (2017). Uji Validitas Pengembangan Tes Untuk Mengukur Kemampuan Pemahaman Relasional Pada Materi Persamaan Kuadrat Siswa Kelas VIII SMP. *Jurnal Penelitian Pembelajaran Matematika Sekolah (JP2MS)*, *1*(1), 60–65. https://doi.org/https://doi.org/10.33369/jp2ms.1.1.60-65

Sektiwulan, A., & Nindiasari, H. (2024). Development of Numeracy Literacy-Based Math Problems in a Personal Context for Middle School Students. *Jurnal Pendidikan Matematika Indonesia*, *9*(2), 174–182. https://doi.org/https://dx.doi.org/10.26737/jpmi.v9i2.5736

Siregar, R., Siagian, M. D., & Syahlan. (2024). Empowering Primary School Students Through Problem-Based Learning: A Path to Literacy and Numeracy Mastery. *Mosharafa: JurnalPendidikan Matematika*, *13*(4), 975–988. https://doi.org/https://doi.org/10.31980/mosharafa.v13i4.2550

Sudarwan, R. E., & Retnawati, H. (2015). Pengembangan Perangkat Assessment Pembelajaran Matematika Pokok Bahasan Geometri Dan Pengukuran SMP/MTs. *Jurnal Riset Pendidikan Matematika*, *2*(2), 251–261. https://doi.org/https://doi.org/10.21831/jrpm.v2i2.7344

Sumarni, S., Prayitno, A. T., Syafari, R., Basir, M. A., Febrianti, D., & Putri, A. (2023). Instrumen Tes Kemampuan Literasi Numerasi Berbasis Budaya Lokal Kabupaten Kuningan Jawa Barat. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *12*(1), 789–801. https://doi.org/10.24127/ajpm.v12i1.6439

Susanto, H., Rinaldi, A., & Novalia, N. (2015). Analisis Validitas Reabilitas Tingkat Kesukaran dan Daya Beda pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika. *Al-Jabar: Jurnal Pendidikan Matematika*, *6*(2), 203–217.

Wardani, D. A., Fathani, A. H., & Alifiani. (2020). Analisis Kemampuan Pemecahan Masalah Matematis Peserta Didik Dalam Menyelesaikan Soal Asesmen Kompetensi Minimum (AKM) Ditinjau Dari Kecerdasan Majemuk. *Jp3: Jurnal Penelitian, Pendidikan, Dan Pembelajaran*, *5*(2), 67–74. https://jim.unisma.ac.id/index.php/jp3/article/view/12486/9669

Wardhani, D. A. P., & Oktiningrum, W. (2022). Pengembangan Soal AKM BermuatanEthnomatematika Dengan Media Canva Untuk Mengukur Kemampuan Literasi Numerasi Siswa Sekolah Dasar. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *11*(4), 3860–3871. https://doi.org/https://doi.org/10.24127/ajpm.v11i4.6241

Yuliyanto, A., Turmudi, T., Syaodih, E., Saputra, D. R., Dharmawan, A., & Pertiwi, C. K. (2021). The Validity and Reliability of Instruments for Measuring Elementary School Students' Early Mathematical Ability. *Kalamatika: Jurnal Pendidikan Matematika*, *6*(2), 127–142. https://doi.org/10.22236/kalamatika.vol6no2.2021pp127-142