AlphaMath
Journal of Mathematics Education, 11(2) November 2025
p-ISSN: 2477-409X, e-ISSN: 2549-9084
Available online: https://jurnalnasional.ump.ac.id/index.php/alphamath/

Development of an Ethnomathematics-Based Test Instrument to Measure and Foster Students' Critical Thinking Skills

Miftahul Fitri[1*], Sri Mariana[2], Istiqomah Nur Ahya[3], Hila Liani[4], Oktami Achni[5], Tri Asthi Wulandari[6], Siti Sinta[7], Arya Tri Agustin[8]
Universitas Islam Sarolangun, Indonesia
miftahulfitri025@gmail.com
http://dx.doi.org/10.30595/alphamath.v11i2.28771

ABSTRACT

Critical thinking is a vital 21st-century skill that must be developed through mathematics learning; however, students' critical thinking performance in Basic Mathematics courses remains low due to conventional assessments that emphasize procedural tasks rather than analytical reasoning. This study aims to develop a valid, reliable, practical, and effective ethnomathematics-based test instrument designed to measure and foster students' critical thinking skills. Employing a Research and Development (R&D) method adapted from Plomp's model, the study involved four stages: preliminary investigation, design, realization/construction, and test, evaluation, and revision. Data were collected using a mathematics test, a practicality questionnaire, and expert validation sheets from 30 third-semester students and four expert validators. Quantitative analysis using SPSS—Aiken's V for content validity, Cronbach's Alpha for reliability, descriptive statistics for practicality, and paired-sample t-tests for pretest–posttest comparison—showed strong psychometric properties: Aiken's V = 0.789 (valid), Cronbach's Alpha = 0.84 (very high reliability), practicality = 88.4%, and improvement from pretest 61.4 to posttest 81.7 (Sig. < 0.05). The study provides empirical evidence that integrating local cultural contexts into mathematics assessments enhances both contextual validity and students' higher-order thinking skills. Therefore, the ethnomathematics-based test instrument is valid, practical, and effective for mathematics evaluation and instruction in higher education, offering a culturally grounded alternative to promote critical thinking in mathematics learning.

Keywords: Critical Thinking, Ethnomathematics, Higher Education, Mathematics Learning, Test Instrument.

Received : November 15, 2025
Revised : November 29, 2025
Accepted : November 30, 2025

Introduction

Mathematics plays a crucial role in education, as it not only develops computational skills but also trains students to think logically, analytically, and critically in solving everyday problems (Sutarto et al., 2022; Yuliana et al., 2023). Therefore, mathematics education at various levels, including higher education, is expected to cultivate students' higher-order thinking skills (HOTS), which are a primary requirement of the 21st century (Amanda et al., 2025; Kamid et al., 2021; Leuwol et al., 2022). In the current era of rapid information growth, critical thinking has become essential for students to make informed decisions and respond wisely to complex challenges. This skill also enables students to adapt to increasingly complex global challenges and develop intellectual autonomy in processing and evaluating information. However, numerous studies indicate that university students' critical thinking skills in solving mathematical problems remain relatively low. Students tend to solve problems by following routine procedures without providing logical reasoning for the steps taken (Driana et al., 2021; Faiziyah & Priyambodho, 2022).

Studies indicate that many university students still tend to solve problems procedurally, focusing on routine calculations without providing logical reasoning or reflective thinking. These observations highlight a clear need for assessment instruments that not only measure students' final answers but also encourage and evaluate their critical thinking processes. Efforts to improve the quality of mathematics education can be undertaken through the development of test instruments capable of both assessing and stimulating students' critical thinking skills. Test instruments serve not only to measure students' learning progress but also to evaluate the success of the teaching and learning process (Ndiung & Jediut, 2020). However, in practice, many lecturers and teachers still use conventional tests that are procedural and do not address higher-order thinking aspects (Ndiung & Jediut, 2020; Ryandi & Santri, 2024). Therefore, it is necessary to develop test instruments that are contextual and relevant to students' lives, enabling them to develop critical thinking skills through meaningful situations (Mauladaniyati et al., 2025).

One approach deemed effective for achieving this goal is ethnomathematics. Ethnomathematics examines the relationship between mathematics and culture and seeks to connect mathematical concepts with practices and values embedded in society (Fitri et al., 2025; Angga et al., 2024; Muhammad et al., 2023). This approach has been shown to help students understand mathematical concepts more concretely because they are presented within a cultural context that is familiar to their daily lives (Dhiya et al., 2025; Wulandari et al., 2024). Realistic learning with an ethnomathematics context is more effective in enhancing mathematical thinking skills than conventional instruction (Martir et al., 2024). The development of ethnomathematics-based problems relevant to local culture, which proved to improve students' problem-solving abilities (Sudrajat, 2025).

Although many studies have explored ethnomathematics in school contexts, research on developing ethnomathematics-based test instruments to assess and train critical thinking skills at the university level remains very limited (Dhiya et al., 2025; Rahmawati & Susilo, 2025). Most previous studies focused on secondary school students and emphasized conceptual understanding rather than critical thinking skills. Furthermore, few studies have linked ethnomathematics to students' own local cultural contexts, despite the fact that local cultural contexts can provide stronger meaning in the mathematical thinking process (Novianti et al., 2023; Siregar, 2025).

Department of Mathematics Education, Universitas Muhammadiyah Purwokerto, Purwokerto, Indonesia
p-ISSN 2477-409X, e-ISSN: 2549-9084

Based on this gap, the present study focuses on developing an ethnomathematics-based test instrument to measure and foster critical thinking skills among students in the Basic Mathematics course. The development process follows four main stages adapted from Plomp's model: preliminary investigation, design, realization/construction, and test, evaluation, and revision, providing a clear overview of how the instrument was created and refined. This study is expected to contribute to the development of mathematics assessment tools that are not only valid and reliable but also meaningful, contextual, and oriented toward strengthening students' critical thinking abilities (Amanda et al., 2025; Andriyani, W. et al., 2021; Qomariyah, 2023).

Methods

Research Design

This study is a development research (Research and Development) with a quantitative approach. The type of development carried out was the development of an ethnomathematics-based test instrument to measure and foster students' critical thinking skills. The development model employed refers to Plomp's model, which consists of four stages: preliminary investigation, design, realization/construction, test, evaluation, and revision. A quantitative research design was chosen to allow systematic measurement of validity, reliability, and practicality, while expert judgment was incorporated to evaluate the instrument's content validity and practicality, ensuring that both numerical data and qualitative assessments contributed to refining and revising the instrument. This research design was chosen because it aligns with the study's objective: to produce an ethnomathematics-based test instrument that is valid, reliable, and practical for measuring and developing students' critical thinking skills in the context of mathematics learning (Amanda et al., 2025; Noer et al., 2021; Ryandi & Santri, 2024).

Participants or Data Sources

The research subjects consisted of 30 third-semester students of the Mathematics Education Study Program who had completed the Basic Mathematics course. Purposive sampling was used, considering the suitability of the students' characteristics for the instrument trial (Mauladaniyati et al., 2025; Nufus & Kusaeri, 2020). The study also involved four expert validators, including two mathematics education lecturers and two experienced mathematics teachers in test instrument development.

Research Instruments

The primary instrument in this study was an essay test consisting of five items. The test was developed based on critical thinking indicators contextualized with ethnomathematics elements in the students' surrounding environment (Martir et al., 2024; Qomariyah, 2023). In addition to the test, the study employed expert validation sheets to assess content and construct validity and a student response questionnaire to evaluate the practicality of the instrument (Putra & Mahmudah, 2021; Wulandari et al., 2024).

Research Procedures

The research procedure followed the four stages of Plomp's model (Noer et al., 2021). In the first stage, preliminary investigation, problems and development needs were identified through literature and curriculum analysis. In the second stage, design, a test blueprint and ethnomathematics-based items aligned with critical thinking indicators were developed. In the third stage, realization/construction, the instrument was created according to the design. In the fourth stage, test, evaluation, and revision, experts assessed content validity and practicality, followed by student trials to examine validity, reliability, difficulty, discrimination, practicality and effectiveness in improving students' critical thinking skills.

Data Collection Techniques

Data were collected through expert evaluations, student practicality questionnaires, and student test results. Expert validators assessed the feasibility and content accuracy of each item to determine the instrument's validity. Practicality data were obtained through student questionnaire responses regarding clarity, usability, and implementation feasibility. Meanwhile, student test results were analyzed to measure the effectiveness of the instrument in improving critical thinking skills.

Data Analysis Techniques

Data were analyzed quantitatively using SPSS 21 to determine the instrument's validity, reliability, difficulty, discrimination, and practicality. The use of SPSS for these analyses is supported by prior studies in development research, which have demonstrated that SPSS provides adequate statistical tools for evaluating psychometric properties and practical aspects of educational instruments (Noer et al., 2021). The instrument was considered feasible for measuring and enhancing students' critical thinking skills if it met the criteria of being valid, reliable, moderately difficult, well-discriminating, and practical (Ryandi & Santri, 2024). Below are the specific analysis conducted:

Validity Test

The validity test aimed to determine the extent to which the instrument accurately measures the intended abilities (Desnita, 2022). In this study, validity was categorized into content validity and empirical validity. Content validity was obtained through the assessment of four expert validators on each item based on content, construct, and language aspects. Content validity was analyzed using the Aiken's V Formula (1) as follows:

$$V = \frac{\sum s}{n(c-1)} \qquad (1)$$

Where: V = content validity coefficient, $s = r - l$, the score given by the validator minus the lowest score on the scale, n = number of validators, c = highest possible score. Aiken's V ranges from 0 to 1, with the instrument considered valid if V>0.5V > 0.5V>0.5. The results of the content validity analysis served as the basis for revising items deemed unsuitable by the validators. Empirical validity was obtained by piloting the instrument with 30 students. Each item score was correlated with the total score using the Pearson Product-Moment correlation Formula (2):

$$r_{xy} = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \qquad (2)$$

An item was considered empirically valid if $r_{calculated} > r_{table}$ at a 5% significance level.

Reliability Test

Reliability measured the consistency of test results under different conditions. Reliability analysis was conducted using Cronbach's Alpha Formula (3):

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_t^2}\right) \qquad (3)$$

Where k is the number of items, σ_i^2 is the variance of each item, and σ_t^2 is the total variance. The instrument was considered reliable if $\alpha > 0{,}6$. Higher alpha values indicate higher reliability.

Difficulty and Discrimination Index Test

The difficulty level of each item was calculated to ensure that the test was neither too easy nor too difficult using the Formula (4):

$$T = \frac{\bar{X}}{SMI} \qquad (4)$$

Where T is the difficulty index, $\bar{X}$ is the average student score per item, and SMI is the ideal maximum score. Items were categorized as: $T < 0{,}30$: difficult; $0{,}31 \leq T \leq 0{,}70$: moderate; $T > 0{,}70$: easy. Items in the moderate category were considered appropriate. The discrimination index measured the extent to which an item distinguished between high- and low-ability students using the Formula (5):

$$DB = \frac{\bar{X}_A - \bar{X}_B}{SMI} \qquad (5)$$

Where DB is the discrimination index test, $\bar{X}_A$ and $\bar{X}_B$ are the average scores of the top and bottom groups, and SMI is the ideal maximum score. The criteria were: $DB < 0{,}20 =$ poor; $0{,}20 \leq DB \leq 0{,}40 =$ fair; $0{,}40 \leq DB \leq 0{,}70 =$ good; $DB \geq 0{,}70 =$ excellent.

Practicality Test

Practicality was measured using a student questionnaire assessing ease of use, clarity of language, and relevance of the ethnomathematics context. A five-point Likert scale with positive and negative statements was used. Practicality percentage was calculated using Formula (6) as:

$$P = \frac{\sum X}{N} \times 100\% \qquad (6)$$

Where P is the practicality percentage, $\sum X$ is the total score obtained, and N is the ideal maximum score. An instrument was considered practical if the average practicality score exceeded 70%. The analyses from all tests were used to determine the feasibility of the ethnomathematics-based test instrument for measuring and fostering students' critical thinking skills. The instrument was deemed feasible if it met the criteria of validity, reliability, moderate difficulty, good discrimination, and practicality.

Effectiveness Test

Effectiveness analysis was conducted to determine whether the instrument could enhance students' critical thinking skills. The improvement in learning outcomes was measured using the N-Gain score, calculated with the Formula (7):

$$N - Gain = \frac{(X_{post} - X_{pre})}{(SMI - X_{pre})} \qquad (7)$$

Where X_{post} is the posttest average score, X_{pre} is the pretest average score, and SMI is the ideal maximum score. The N-gain category was classified as follows: $g > 0.70 =$ high, $0.30 \leq g \leq 0.70 =$ moderate, $g < 0.30 =$ low. The test instrument was considered effective if the N-Gain score achieved at least the moderate category.

Result and Discussions

Development Results

The following is a detailed description of the development process, testing results, and analysis of the final product of the developed test instrument

Preliminary Stage

The preliminary stage began with determining the research location and subjects. This study was conducted in the Mathematics Education Study Program at Universitas

Islam Sarolangun with 30 third-semester students as participants. This stage identified the need for an ethnomathematics-based test instrument to measure and foster students' critical thinking skills. Interviews and observations revealed that while students understood basic mathematical concepts, they struggled to relate them to local cultural contexts such as traditional patterns, regional architecture, and community calculation practices. The absence of a test integrating ethnomathematics with critical thinking indicators highlighted the need to develop a contextual essay-based instrument.

Self-Evaluation Stage

This stage consists of two parts: analysis and design.

Analysis

The analysis covered; Curriculum analysis, to ensure the instrument aligns with basic mathematics learning outcomes and emphasizes critical thinking in real-life contexts (Wulandari et al., 2024). Student characteristic analysis, based on interviews and tasks, students had adequate mathematical knowledge but struggled with reasoning and argumentation, key components of critical thinking (Ryandi & Santri, 2024). Material analysis, identifying mathematical concepts and relevant cultural contexts to construct questions, focusing on basic mathematics topics easily connected to cultural elements (Angga et al., 2024).

Design

At this stage, the design included: the instrument blueprint, five ethnomathematics-based essay items, answer keys, and scoring rubrics. Each question was constructed based on critical thinking indicators by Facione (2015): interpretation, analysis, evaluation, inference, and explanation. The initial product was called Prototype I, and Figure 1 shows the layout of the ethnomathematics-based test instrument, which was subsequently validated by the experts:

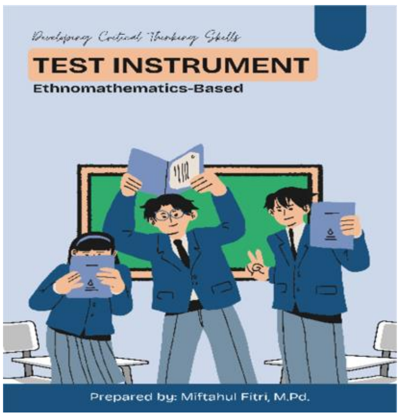


Figure 1. Design of ethnomathematics-based test instrument

Prototyping Stage

This stage included three steps: expert review, one-to-one, and small group evaluation. Expert Review; Prototype I was validated by four experts—two mathematics education lecturers and two experienced mathematics teachers. They evaluated the content, construct, and language aspects. Table 1 explains the results.

Table 1. Assesment Result by the Validator

Assessment Aspect	Validity Coefficient
Validator I	Feasible with minor revisions
Validator II	Feasible without revision
Validator III	Feasible without revision
Validator IV	Feasible without revision

Revisions at this stage involved: clarifying ambiguous items, improving sentence structure, and adjusting language to align with students' comprehension level. These changes produced a refined Prototype I.

One to One; Prototype I was tested on three students representing high, medium, and low ability levels. Students completed the test and filled out a readability response questionnaire. Based on feedback, items with unclear wording were revised, some instructions simplified, and cultural references refined. These modifications produced Prototype II, with improved clarity and accessibility..

Small Group; Prototype II was tested on six students with varying abilities. They completed the test and filled out a response questionnaire on instruction clarity, readability, and cultural context relevance. Revisions from this stage included further refinement of ambiguous items, adjustment of cultural contexts for relevance, and improvement of overall readability, producing Prototype III, which was ready for field testing.

Field Test Stage

Prototype III was administered to 30 third-semester students. The test was conducted in one meeting, with 60 minutes allocated for five essay questions. Data from the field test were analyzed for empirical validity, reliability, difficulty level, discrimination power, practicality, and effectiveness (Darma, 2021).

Results of the Ethnomathematics-Based Test Instrument

Content Validity Analysis

Content validity was measured using Aiken's V formula to assess item alignment with the intended indicators.

Table 2. Results Content Validity Analysis

No	Assessment Aspect	Validity Coefficient	Criteria
1	Content	0.782	Valid
2	Construct	0.808	Valid
3	Language	0.776	Valid
	Average	0.789	Valid

In Table 2, the average validity value of 0.789 indicates that the instrument is valid and suitable for testing. During the validation process, the experts suggested minor revisions in wording and clarity to ensure the questions were more communicative and culturally relevant. These suggestions were incorporated, enhancing both the clarity and contextual appropriateness of the items.

Emperical Validity Analysis

Empirical validity was tested with 30 students using the Product Moment correlation formula. The r_{table} value at N = 30 and 5% significance was 0,361.

Table 3. Results Emperical Validity Analysis

No	$r_{calculated}$	r_{table}	Criteria
1	0.721	0.361	Valid
2	0.667	0.361	Valid
3	0.548	0.361	Valid
4	0.764	0.361	Valid
5	0.624	0.361	Valid

Based on Table 3, all five items had $r_{calculated} > r_{table}$, meaning all were empirically valid and appropriate for further research. During item development, some items were refined based on student responses to better reflect critical thinking processes, including reasoning, reflection, and explanation.

Student Response Analysis

The practicality of the instrument was analyzed using response questionnaires from the one-to-one and small group stages, which consisted of 15 statement items. Overall, student responses was 76.2% (positive), indicating that the ethnomathematics-based test instrument was practical.

Table 4. Summary of Student Response Questionnaire

Test Stage	Statement with the Highest Response	Percentage	Statement with the Lowest Response	Percentage	Average	Category
One to One	No. 14 & 15 (the use of communicative	85%	No. 9 (the time allocation for	45%	72,3	Positive

	language and the depiction of local culture)		completing the test items)			
Small Group	No. 14	94%	No. 3 (the relevance of the questions to daily life)	65%	80,1	Positive
		Avarage			76,2	Positive

From the Table 4, it can be concluded that the student response questionnaire meets positive criteria, indicating that the ethnomathematics-based test instrument is practical and well-accepted by students. Students reported that the questions were engaging, understandable, and relevant to real-life contexts, which demonstrates qualitative evidence of practicality.

Realibility Analysis

Using Cronbach's Alpha, the reliability coefficient was α = 0.84, categorized as very high**,** indicating strong internal consistency.

Table 5. Dificulty Level Analysis

Question Number	Difficulty Index	Category
1	0.72	Easy
2	0.64	moderate
3	0.58	moderate
4	0.69	moderate
5	0.63	moderate
Rata-Rata	0.65	moderate

Difficulty Level and Discrimination Power

Most items were of moderate difficulty (0.65 average) and had fair discrimination power (0.30 average)**,** making them effective for distinguishing between high and low critical-thinking students.

Table 6. Discrimination Power

Question Number	Discrimination	Category
1	0.28	Fair
2	0.31	Fair
3	0.34	Fair
4	0.27	Fair
5	0.32	Fair
Rata-Rata	0.30	Fair

Based on Table 5, all five test items fall within the moderate category, making them ideal for measuring critical thinking skills. Based on Table 6, all five items are categorized as fair, with an average discrimination index of 0.30 (fair). This indicates that the developed test instrument is adequate for distinguishing between students with high and low levels of critical thinking ability.

Practicality and Effectiveness Analysis

The practicality and effectiveness of the ethnomathematics-based test instrument were evaluated through limited (10 students) and field trials (30 students). In Tabel 7, student response questionnaires indicated positive results, with practicality percentages of 84.6% and 88.4%, categorized as practical and highly practical.

Table 7. Results of the Test Instrument Practicality Test

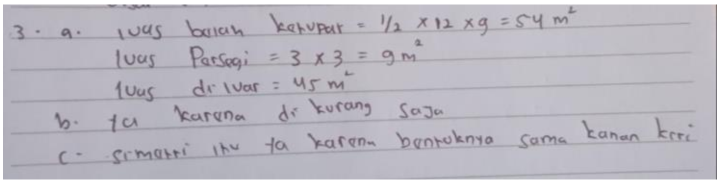
Test Stage	Number of Respondents	Average Score	Percentage (%)	Criteria
Small-Scale Trial	10	42.3	84.6	Practical
Large-Scale Trial	30	126.7	88.4	Highly Practical

Student response questionnaires indicated positive results, and students found the questions engaging, clear, and relevant to their daily lives. Furthermore, the effectiveness analysis demonstrated an improvement in students' critical thinking skills, as indicated by the increase in the average score from 61.4 (pretest) to 81.7 (posttest). Based on the N-Gain calculation (0.53), which falls into the moderate category, the ethnomathematics-based test instrument is considered effective in enhancing students' critical thinking skills. It should be noted that these scores reflect students' learning outcomes, and not the intrinsic effectiveness of the instrument itself. The qualitative feedback from students and observations during trials also confirmed that the instrument encouraged reasoning, reflection, and logical explanations, demonstrating its ability to stimulate critical thinking processes in meaningful contexts.

Table 8. Results of the Test Instrument Effectiveness Test

Test	Average Score	N	Criteria
Pretest	61,4	30	Fair
Posttest	81,7	30	Good

Based on Table 8, in addition to the improvement in pretest and posttest scores, the instrument's effectiveness is also reflected in the qualitative differences between the responses of students with low and high levels of critical thinking on ethnomathematics-based questions.



3. a. luas belah ketupat = 1/2 x 12 x 9 = 54 m²
luas Persegi = 3 x 3 = 9 m²
luas di luar = 45 m²
b. ya karena di kurang saja
c. simetri itu ya karena bentuknya sama kanan kiri

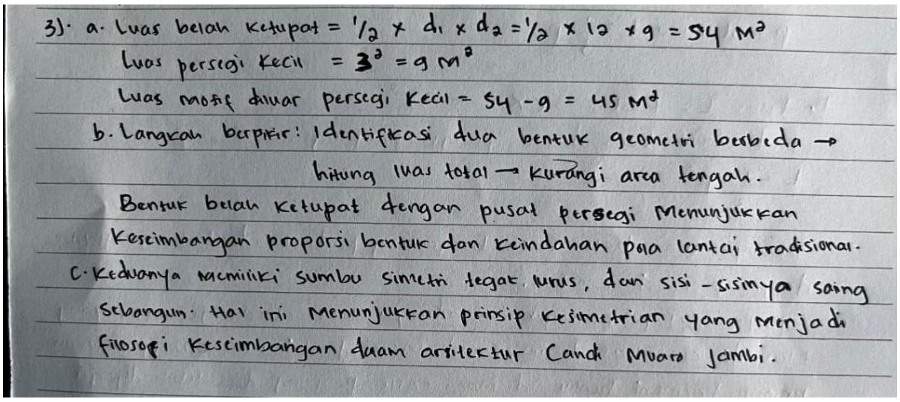
Figure 2. Answers of students with low critical thinking ability

Translate Figure 2 in English:

a. Area of rhombus $= \frac{1}{2} \times 12 \times 9 = 54\ m^2$
Area of square $= 3 \times 3 = 9\ m^2$
Area outside the square $= 45\ m^2$

b. Not effective because the result is less (incomplete)

c. Symmetry exists because the shape is the same on the right and left sides



3). a. Luas belah ketupat = 1/2 × d1 × d2 = 1/2 × 12 × 9 = 54 m²
Luas persegi kecil = 3² = 9 m²
Luas motif diluar persegi kecil = 54 − 9 = 45 m²
b. Langkah berpikir: identifikasi dua bentuk geometri berbeda → hitung luas total → kurangi area tengah.
Bentuk belah ketupat dengan pusat persegi menunjukkan keseimbangan proporsi bentuk dan keindahan pola lantai tradisional.
c. Keduanya memiliki sumbu simetri tegak lurus, dan sisi-sisinya saling sebangun. Hal ini menunjukkan prinsip kesimetrian yang menjadi filosofi keseimbangan dalam arsitektur Candi Muaro Jambi.

Figure 3. Answers of students with high critical thinking ability

Translate Figure 3 in English:

a. Area of rhombus $= \frac{1}{2} \times d_1 \times d_2 = \frac{1}{2} \times 12 \times 9 = 54\ m^2$
Area of square $= 3^2 = 9\ m^2$
Area outside the square $54 - 9 = 45\ m^2$

b. Reasoning steps: identify two different geometric shapes, calculate the total area, and subtract the central area. The rhombus shape with a square in the center shows the balance of proportion and harmony in traditional floor patterns of temples

c. Both shapes have symmetrical sides and equal diagonal alignment. This illustrates the principle of symmetry, which reflects the philosophy of balance in the architecture of Muaro Jambi Temple

In Figure 2, a student with low critical thinking ability correctly calculated the rhombus area (54 m²) and subtracted the small square area (resulting in 45 m²). However, the explanation remained procedural and limited to geometric descriptions such as *"yes, just subtract it"* or *"symmetry means the same left and right."* This shows that

Department of Mathematics Education, Universitas Muhammadiyah Purwokerto, Purwokerto, Indonesia
p-ISSN 2477-409X, e-ISSN: 2549-9084

the student failed to connect the results with symmetry, balance, or cultural meaning in the temple ornament context. Conversely, in Figure 3, a student with high critical thinking ability demonstrated analytical and reflective reasoning by linking geometric balance and the cultural philosophy of harmony in Muaro Jambi Temple architecture. Statements like *"the rhombus shape with a central square represents proportional balance and traditional floor pattern aesthetics"* illustrate deep integration of mathematical and cultural understanding. This comparison reinforces that ethnomathematics-based test instruments improve not only numerical accuracy but also the ability to think critically about mathematical and cultural interconnections.

Discussion

The development of this instrument followed the Plomp model, consisting of the preliminary, self-evaluation, prototyping, and field test stages. The final product was an ethnomathematics-based test instrument comprising five essay items embedded in local cultural contexts. Each item reflected critical thinking indicators such as analysis, evaluation, and inference (Facione, 2015). During development, students demonstrated better understanding when mathematical problems were presented through familiar cultural narratives, such as regional architectural patterns.

The content validity value (0.789) and empirical validity results confirmed that all items were valid. This supports (Kurniawan & Syafriani, 2021), who stated that an Aiken's V ≥ 0.7 indicates high validity. The student response questionnaire showed an average of 76.2% (positive). Students found the questions easy to understand, communicative, and culturally relevant—effective for developing critical thinking skills. The reliability coefficient (0.84) showed excellent internal consistency, indicating stable and consistent measurements. Reliability as a key indicator of measurement quality in education (Arsi, 2021). Analysis revealed that all items were of moderate difficulty and fair discrimination. This means the questions were neither too easy nor too difficult and effectively distinguished between students with high and low levels of critical thinking. Moderately difficult items are ideal for assessing higher-order thinking (Saputri & Larasati, 2023). The practicality test yielded an average score of 88.4%, categorized as highly practical. Students reported that the questions were engaging, comprehensible, and relevant to real-life contexts. Effectiveness was assessed using the N-Gain score, which was 0.53, falling into the moderate category. This indicates that the instrument is effective in improving students' critical thinking skills, with average scores increasing from 61.4 (pretest) to 81.7 (posttest). These findings are consistent with the literature on the development of mathematics HOTS assessment instruments, which report instruments that are valid, reliable, and contextually relevant (Ryandi & Santri, 2024).

The findings of this study highlight the potential of ethnomathematics-based assessment to meaningfully improve students' critical thinking skills in higher education. This suggests that integrating cultural contexts into assessment can support more relevant and meaningful learning. Therefore, educators are encouraged to adopt similar culturally grounded assessment approaches. Future research is recommended to involve broader and more diverse samples and to explore technology-based assessment systems to increase scalability.

Conclusion

This study developed an ethnomathematics-based test instrument that is valid, reliable, practical, and effective for measuring and enhancing students' critical thinking skills in Basic Mathematics. The instrument showed strong validity (Aiken's V = 0.789), high reliability (α = 0.84), and appropriate difficulty and discrimination levels. Practicality and effectiveness were confirmed through positive student responses (88.4%) and a significant improvement in scores from 61.4 to 81.7. These results extend the ethnomathematics approach into assessment, demonstrating that integrating local culture enhances contextual validity and supports meaningful learning. Although limited to one program, this research lays a foundation for developing contextual, character-based mathematics assessments, with future studies recommended to explore broader cultural contexts and technology-based applications.

This study has several limitations that should be acknowledged. First, the trial was conducted with a relatively small sample from a single university, which may limit the generalizability of the findings. Second, the ethnomathematics context used in the instrument was specific to a certain local culture, so its suitability in different cultural settings may vary. Third, the effectiveness measure focused only on short-term improvement through pretest–posttest comparison, without examining long-term impact.

Acknowledgement

The researcher sincerely thanks to Universitas Islam Sarolangun, especially the Mathematics Education Department, for their full support throughout this research. Gratitude is also extended to the lecturers for their valuable suggestions and to the Mathematics Education students for their active participation during the instrument trials.

Author's Declaration

Author Contribution	:	Author 1: Conceptualization, Methodology, Data Curation, Writing - Original Draft Author 2: Validation, Formal Analysis, Writing – Review & Editing Author 3: Instrument Development, Data Collection, Visualization Author 4: Investigation, Data Collection, Project Administration Author 5: Formal Analysis, Writing – Review & Editing, Resources Author 6: Software, Statistical Analysis, Validation Author 7: Literature Review, Documentation, Writing – Editing Author 8: Supervision, Conceptualization, Final Approval of the Manuscript
Funding Statement	:	This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Conflict of Interest	:	The authors declare no conflict of interest.
Additional Information	:	-

References

Amanda, N. L., Armanto, D., & Rajaguguk, W. (2025). *HOTS Test Development Based on Realistic Mathematics Learning For Grade V Elementary School*. *9*(1), 1–11. https://doi.org/10.15575/jp.v9i1.328

Andriyani, W., Suparman, & Ilato, R. (2021). Developing Critical Thinking of Students with Hearing Impairment for Computational Thinking in Mathematics with E Module Design. *Psychology and Education Journal*, *58*(1), 4613–4621. https://doi.org/10.17762/pae.v58i1.1578

Angga, R., Tuğba, P., & Yelken, Y. (2024). Effectiveness of ethnomathematics - based learning on students ' mathematical literacy : a meta - analysis study. *Discover Education*, *3*(1), 202-212. https://doi.org/10.1007/s44217-024-00309-1

Arsi, A. (2021). Realibilitas Instrumen Dengan Menggunakan Spss. *Validitas Realibilitas Instrumen Dengan Menggunakan Spss*, 1–8.

Darma, B. (2021). Statistika Penelitian Menggunakan SPSS (Uji Validitas, Uji Reliabilitas, Regresi Linier Sederhana, Regresi Linier Berganda, Uji t, Uji F, R2). *Guepedia*. https://books.google.co.id/books?id=acpLEAAAQBAJ

Desnita, D. (2022). Validity and Reliability of Critical Thinking Instruments to Measure the Effectiveness of Context-Based Physics E-Module on Wave Materials. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, *8*(1), 57–64. https://doi.org/10.21009/1.08106

Dhiya, R., Iffah, L., Subanti, S., Usodo, B., & Nurhasanah, F. (2025). *Systematic literature review : Ethnomathematics research in Indonesia*. *10*(1), 28–40. https://doi.org/10.23917/jramathedu.v10i1.5621

Driana, E., Susilowati, A., Ernawati, E., & Ghani, A. R. A. (2021). *Assessing students' higher-order thinking skills: Knowledge and practices of chemistry teachers in vocational senior secondary schools*. *Jurnal Pendidikan Teknologi Dan Kejuruan, 27*(1), 37–47.

https://doi.org/10.21831/jptk.v27i1.32882

Facione, P. a. (2015). Critical Thinking : What It Is and Why It Counts. (ISBN 13: 978-1-891557-07-1). *Insight assessment*, *1*(1), 1-23. https://www.insightassessment.com/article/critical-thinking-what-it-is-and-why-it-counts/

Faiziyah, N., & Priyambodho, B. L. (2022). Analisis kemampuan berpikir kritis dalam menyelesaikan soal HOTS ditinjau dari metakognisi siswa. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *11*(4), 2823–2835. https://doi.org/10.24127/ajpm.v11i4.5918

Fitri, M., Mujahidawati, & Falani, I. (2025). Pengembangan E-Modul Etnomatematika 3d Terintegrasi Virtual Reality Berbasis Tpack Untuk Meningkatkan Kemampuan Berpikir Kritis Siswa. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *14*(2), 368-380. https://doi.org/10.24127/ajpm.v14i2.10191

Kamid, K., Saputri, R., & Hariyadi, B. (2021). Pengembangan Soal Higher Order Thinking Skills Berbasis Budaya Jambi. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, *5*(2), 1793–1806. https://doi.org/10.31004/cendekia.v5i2.678

Kurniawan, R., & Syafriani. (2021). The validity of e-module based on guided inquiry integrated ethnoscience in high school physics learning to improve students' critical thinking. *Journal of Physics: Conference Series*, *1876*(1), 1-7. https://doi.org/10.1088/1742-6596/1876/1/012067

Leuwol, F. S., Busnawir, M. S., Saryanto, S. P. T., Retnaningsih, R., Amalia, R., Sembiring, T. B., Mardikawati, B., Bambang Sucipto, M. M., Doho, Y. D. B., & S Phil, M. M. (2022). *Kemampuan Berpikir Tingkat Rendah (LOTS) VS Kemampuan Berpikir Tingkat Tinggi (HOTS)*. Indramayu: Penerbit Adab

Martir, L., Beku, V. Y., Wewe, M., & Dhiu, L. M. (2024). Upaya Meningkatkan Kemampuan Berpikir Kritis Matematis Melalui Pendekatan Etnomatematika. *Polinomial: Jurnal Pendidikan Matematika*, *3*(2), 66–74. https://doi.org/10.56916/jp.v3i2.914

Mauladaniyati, R., Purnomo, H., & Wijayanti, P. (2025). Meta-Analysis of the Impact of Ethnomathematics-Based Digital Learning on Students ' Mathematical Understanding. *Journal of Emerging Technologies in Ethnomathematics, 1*(1), 55–66. https://doi.org/10.26740/jetie.v1i1.41039

Muhammad, I., Marchy, F., Naser, A. D. M., & Turmudi. (2023). Analisis Bibliometrik: Tren Penelitian Etnomatematika dalam Pembelajaran Matematika Di Indonesia (2017–2022). *JIPM (Jurnal Ilmiah Pendidikan Matematika)*, *11*(2), 267–279. https://doi.org/10.25273/jipm.v11i2.14085

Nasution, E. (2024). *Pengembangan perangkat pembelajaran beriorentasi pendekatan matematika realistik berbasis etnomatematika untuk meningkatkan kemampuan pemecahan masalah matematis siswa*. (Doctoral dissertation UIN Syekh Ali Hasan Ahmad Addary Padangsidimpuan. Retrieved from https://etd.uinsyahada.ac.id/12202

Ndiung, S., & Jediut, M. (2020). Pengembangan instrumen tes hasil belajar matematika peserta didik sekolah dasar berorientasi pada berpikir tingkat tinggi. *Premiere Educandum: Jurnal Pendidikan Dasar Dan Pembelajaran*, *10*(1), 94-112 https://doi.org/10.25273/pe.v10i1.6274

Noer, A. M., Putri, E. N., Rery, R. U., Anwar, L., & Tarawi, O. (2021). The e-module development of reaction rate based on guided inquiry as independent teaching materials. *Journal of Physics: Conference Series*, *1806*(1), 1-7. https://doi.org/10.1088/1742-6596/1806/1/012176

Novianti, N., Zaiyar, M. Z., Khaulah, S., Fitri, H., & Jannah, R. (2023). Pengembangan E-Modul Berbasis Problem Based Learning Terhadap Kemampuan Berfikir Kritis Siswa. *JISIP (Jurnal Ilmu Sosial Dan Pendidikan)*, *7*(3), 2369–2375. https://doi.org/10.58258/jisip.v7i3.5370

Nufus, H., & Kusaeri, A. (2020). Analisis tingkat kemampuan berpikir kritis siswa dalam memecahkan masalah geometri. *Jurnal Pendidikan Matematika Indonesia*, *5*(2), 49–55. https://journal.stkipsingkawang.ac.id/index.php/JPMI/article/view/1812

Putra, E. C. S., & Mahmudah, F. N. (2021). The implementation of ethnomathematics based-learning for students. *SJME (Supremum Journal of Mathematics Education)*, *5*(2), 162-169. https://doi.org/10.35706/sjme.v5i2.4827

Qomariyah, R. (2023). Pengembangan Instrumen Tes Berbasis Etnomatematika Untuk Melatih Kemampuan Pemecahan Masalah Siswa Sekolah Menengah Pertama. *Konstanta : Jurnal Matematika dan Ilmu Pengetahuan Alam*, *1*(2), 181-195. https://doi.org/10.59581/konstanta-widyakarya.v1i2.826

Rahmawati, N. D., & Susilo, G. (2025). Kemampuan berpikir kritis mahasiswa pada pembelajaran problem based learning bernuansa etnomatematika. *De Fermat: Jurnal Pendidikan Matematika*, *8*(1), 397–404. https://doi.org/10.36277/defermat.v8i1.2311

Ryandi, R. B., & Santri, D. D. (2024). *Literature Review On Development Of Mathematics Higher Order Thinking Skills (HOTS) Assessment Instruments*. *Journal of Curriculum and Pedagogic Studies (JCPS)*, *3*(2), 57-69. https://doi.org/10.30631/jcps.v3i2.2908

Saputri, H. A. S., & Larasati, N. J. (2023). Analisis Instrumen Assesmen: Validitas, Reliabilitas, Tingkat Kesukaran Dan Daya Beda Butir Soal. *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, *9*(5), 2986–2995. https://doi.org/10.36989/didaktik.v9i5.2268

Siregar, T. (2025). *Integrasi etnomatematika dengan kearifan budaya lokal*. Cirebon: Goresan Pena.

Sudrajat, S. (2025). Pengembangan Soal Three-Tier Multiple Choice Berbasis Etnomatematika Untuk Mengukur Kemampuan Literasi Matematika Siswa. *ELIPS: Jurnal Pendidikan Matematika*, *6*(1), 57–69. https://doi.org/10.47650/elips.v6i1.1758

Sutarto, Muzaki, A., Hastuti, I. D., Fujiaturrahman, S., & Untu, Z. (2022). Development of an Ethnomathematics-Based e-Module to Improve Students' Metacognitive

Ability in 3D Geometry Topic. *International Journal of Interactive Mobile Technologies*, *16*(3), 32–46. https://doi.org/10.3991/IJIM.V16I03.24949

Wulandari, D. U., Mariana, N., & Amien, M. S. (2024). *Integration of Ethnomathematics Teaching Materials in Mathematics Learning in Elementary School*. *5*(1), 204–218. https://doi.org/10.46245/ijorer.v5i1.542

Yuliana, N., Ansori, H., & Juhairiah, J. (2023). Pengembangan Soal Pemecahan Masalah Matematika Berbasis Computational Thinking Dengan Konteks Etnomatematika. *EDU-MAT: Jurnal Pendidikan Matematika*, *13*(1), 115–127. https://doi.org/10.20527/edumat.v13i1.20250