

Assessing Geometric Proof Skills: Instrument Validation Through Reliability and Factor Analysis

Lathiful Anwar
Universitas Negeri Malang

Cholis Sa'dijah
Universitas Negeri Malang

Tomi Listiawan
Universitas Negeri Malang

Received: March 2025 || Accepted: August 2026

© 2026 Mathematics Education Research Group of Australasia, Inc.

From a mathematics teacher education perspective, identifying the component processes involved in geometric proof reasoning may help prospective teachers interpret learners' proof attempts and plan instruction. This study examined the internal consistency and empirical structure of a 25-item multiple-choice instrument for assessing proof-related reasoning in geometry. Responses from 275 Indonesian preservice mathematics teachers were analysed using item-total correlations, Cronbach's alpha, and an exploratory analysis of the instrument's component structure. The overall internal consistency was high (Cronbach's $\alpha = 0.892$), although four items had corrected item-total correlations below 0.30. The reported component analysis yielded a six-component solution accounting for 56.63% of the total variance. The components concern deductive relationships, linking premises, identifying relevant statements, recognising circular arguments, and reasoning with particular and universal propositions. These findings provide preliminary evidence of the instrument's usefulness for investigating proof-related reasoning. However, weak items, cross-loadings, and the need for further psychometric confirmation mean that its subscale interpretations should be considered provisional. In mathematics teacher education, responses to the instrument may help teacher educators design activities in which prospective teachers analyse reasoning errors and consider appropriate instructional responses.

Keywords • geometric proof construction • instrument validation • reliability • exploratory factor analysis • deductive reasoning • preservice teachers

Introduction

Mathematical proofs are crucial for developing mathematical knowledge and reasoning. In geometry education, proofs are not only a formal requirement but also a process through which students learn to justify mathematical relationships and develop logical arguments (Jablonski & Ludwig, 2023). Constructing geometric proofs requires students to coordinate multiple reasoning processes, including identifying relevant premises, making logical inferences, and forming valid conclusions (Ramírez-Uclés, & Ruiz-Hidalgo, 2022). These processes are essential components of mathematical reasoning and widely recognised as fundamental elements of mathematics learning (Fujita et al., 2018; Jones et al., 2015; Stylianides et al., 2024).

In mathematics education, it is important to distinguish between related and conceptually distinct constructs. Logical reasoning refers to the ability to derive conclusions from given premises applying valid rules of inference (Stylianides et al, 2024). In contrast, critical thinking involves evaluating the validity of arguments, questioning assumptions, and identifying inconsistencies in reasoning (King, 2019)). Contrastingly, proof construction skills specifically refer to the ability to develop structured mathematical arguments that justify mathematical statements based on definitions, axioms, and previously established theorems (Lee, 2026). Although these constructs are related, proof construction requires the integration of logical reasoning with mathematical knowledge and argumentation skills (Anwar et al., 2023; Miyazaki & Fujita, 2015; Yang & Lin, 2008).

Previous research has demonstrated that constructing formal proofs presents a significant challenge to students (Cirillo & Hummer, 2021; Stylianides et al., 2024). Examples of such challenges include reasoning to support the proof, identification of relevant premises, logical connectors of statements, and avoidance of fallacies, such as circular reasoning. Students also have difficulty understanding the structure of the proof and the interdependence of the components of mathematical arguments (Cirillo & Hummer, 2021; Miyazaki & Fujita, 2015). Consequently, proof construction is one of the more difficult aspects of mathematics education, and the development of students' proof construction abilities is a persistent problem across the discipline.

This situation is especially important for preservice mathematics teachers, who are required to support students' development of proof-based reasoning in secondary school mathematics. Effective teaching of proof requires teachers to possess not only knowledge of mathematical proof but also knowledge of how students construct, interpret, and justify proofs and how this reasoning can be supported through instruction (Buchbinder & McCrone, 2020; Leung & Lee, 2017). Consequently, preservice teachers would benefit from knowledge about geometric proofs and the proof-construction process to explain and justify students' reasoning (Leung & Lee, 2017), as that knowledge provides a foundation for developing the mathematical and pedagogical knowledge required to teach reasoning and proof effectively (Lee, 2026). Hence, the proof-construction skills of preservice teachers have direct implications for the provision of preservice teacher education.

Although proof construction is a critical element of mathematics education, few studies have developed and validated assessment tools for constructing geometric proofs. Numerous researchers have examined the reasoning processes of students performing proof tasks; however, most of these studies have either been based on qualitative analysis or task-centric observations, rather than using a standardised assessment tool (Anwar et al., 2021; Fujita & Jones, 2014; Yang & Lin, 2008). Therefore, the evidence of proof construction indicates the potential for the development of standardised measurement tools capable of assessing proof construction from multiple perspectives. Assessments developed to measure proof construction have often only focused on dimensions of proving, including general problem-solving processes, deductive reasoning, and proof comprehension (Makridis, 2022), rather than investigating the full complexity of proof construction (Stein et al., 2026).

Additionally, many of the existing assessments are underdeveloped psychometrically, especially regarding proof-related reasoning skills, which indicates a potential need to examine the mechanisms that proof-related reasoning skills attempt to measure. These gaps reinforce the need to create and refine assessments that focus on the construction of geometric proofs. In response to this need, the current study focuses on constructing a reliable and valid assessment tool to measure the geometric proof construction skills of preservice mathematics teachers. In this assessment, it is expected that multiple reasoning components of proof construction, such as logical reasoning, premise identification, and drawing valid conclusions from a geometric proof, will be assessed. This study aims to develop and validate a tool to assess students' geometric-proof construction skills toward aiding research and instructional enhancement in mathematics by analysing the psychometric properties of the instrument. Specifically, this study evaluates the reliability, item validity, and underlying factor structure of the instrument developed using exploratory factor analysis (EFA). Through this validation process, this study aims to contribute to the development of more robust measurement tools that can be employed to evaluate and support the development of proof-related reasoning skills among preservice mathematics teachers.

Methods

Research Design

A quantitative instrument-evaluation design (Cohen et al., 2007; Creswell, 2014) was used to examine the internal consistency, item functioning, and empirical structure of an assessment of geometric proof-

related reasoning among preservice mathematics teachers. Corrected item–total correlations and Cronbach’s alpha were calculated in SPSS version 29. The component solution was explored using the extraction and rotation procedures reported with Tables 2 and 3. The results are interpreted as preliminary structural evidence rather than definitive confirmation of six independent latent traits.

Participants

The participants were 275 undergraduate mathematics education students from seven Indonesian universities who were preparing to become secondary-school mathematics teachers. Lecturers at participating universities distributed the online instrument to eligible students. Eligibility required enrolment in an undergraduate mathematics education programme, completion of at least one geometry course, and voluntary agreement to participate. This ensured that participants had the knowledge required to deal with tasks involving the construction of geometric proofs.

Stratified random sampling (Creswell, 2014) was used so that each university involved in the study was represented. The rationale for selecting participants from different universities was based on the need to achieve diversity in the sample. Recruitment through participating lecturers meant that the sample was institutionally diverse but was not a probability sample of Indonesian preservice teachers. The participants included 198 women (72%) and 77 men (28%); 96 were first-year, 121 second-year, and 58 third-year students. All had completed at least one course involving geometry. The inclusion of students from seven institutions broadened the range of participating university contexts but does not, by itself, establish national representativeness or generalisability.

Instrument: Geometric Proof Reasoning Assessment

The study used a 25-item multiple-choice assessment of reasoning involved in constructing geometric proofs (hereafter, the Geometric Proof Reasoning Assessment [GPRA]—descriptive title introduced for this manuscript). Instrument development was informed by literature on geometric proof, proof comprehension, deductive reasoning (Makridis, 2022), mathematical logic (Anwar et al., 2021, 2023; Jones et al., 2015; Miyazaki et al., 2017), and task design (Komatsu & Jones, 2019). The items were intended to promote cognitive engagement and elicit recognition and selection of logically appropriate statements and steps in a geometric argument (Hohol & Miłkowski, 2019). Because participants selected answers rather than independently writing complete proofs, the instrument measures selected aspects of proof-related reasoning, not observed proficiency in composing full proofs.

Items were organised conceptually around six intended aspects: drawing deductive inferences, interpreting premises and conclusions, distinguishing particular from universal claims, identifying circular reasoning, chaining conditional statements, and forming a coherent argument. Participants examined geometrical statements and selected answers concerning logical relationships and conclusions. The items and response options are reproduced in the Appendix. The empirically derived components should not be presumed to correspond exactly to these initially intended aspects.

A preliminary pilot with 32 preservice mathematics teachers at a university not included in the main sample examined item wording, readability, administration time, and comprehension of instructions. Feedback informed wording revisions before administration to the principal sample. The final assessment items are reproduced in the Appendix.

Data Collection

Data were collected during a one-month period using a Google Forms link distributed through lecturers at the participating institutions. Eligible students completed the assessment online. Participation was voluntary; the consent and confidentiality procedures are described under Ethical Considerations.

Data Analysis

The data collected were analysed in three stages: validity testing, reliability analysis, and exploratory factor analysis (EFA).

Validity Testing

Corrected item–total correlations were used to examine how strongly each item related to the total score calculated from the remaining items. Values of 0.30 or greater were treated as a screening guideline for item functioning, not as a stand-alone demonstration of construct validity (Field, 2017). Items below this guideline were flagged for additional qualitative and psychometric review.

Reliability Analysis

Internal consistency of the 25-item total score was examined using Cronbach's alpha in SPSS 29. An alpha of 0.70 is often used as a pragmatic screening benchmark, although interpretation depends on item content and score use. The observed alpha of 0.892 indicated high internal consistency in this sample. Alpha-if-item-deleted values and corrected item–total correlations were also examined to identify potentially weak items.

Exploratory Factor Analysis

An exploratory component analysis was reported to investigate patterns of covariation among the 25 items. Sampling adequacy was evaluated using the Kaiser–Meyer–Olkin (KMO) statistic and Bartlett's test of sphericity (Hair et al., 2013). The present tables explicitly label the extraction method as principal component analysis (PCA); accordingly, the reported numerical results are described as a PCA component solution, rather than being presented as principal axis factoring results. Confirmation of the original SPSS extraction settings is required before a claim about common-factor EFA can be made.

The reported analysis used Varimax rotation with Kaiser (1974) normalisation and retained six components with initial eigenvalues greater than one. Loadings of at least 0.40 were treated as salient for descriptive interpretation, with cross-loadings considered rather than suppressed. Because the instrument comprises multiple-choice items, future validation should examine whether an analysis based on item-appropriate correlations, parallel analysis, and independent-sample confirmation supports the proposed structure.

Ethical Considerations

Ethical approval for this study was obtained from all participating institutions. Participation in this study was voluntary. Before completing the questionnaire, the participants were provided with detailed information regarding the purpose of the study, the voluntary nature of their participation, the confidentiality of their responses, and their right to withdraw from the study at any time without penalty. Written informed consent was obtained from all participants prior to data collection. All responses were anonymised before data analysis was conducted to ensure confidentiality and to protect the participants' privacy.

Result

The results describe the item statistics, total-score internal consistency, and exploratory component structure of the 25-item assessment. These analyses provide evidence about score functioning in the present sample (Shrestha, 2021), while leaving questions about external validity and stability of the component structure for future study.

The reported extraction and rotation outputs in Tables 2 and 3 identify principal component analysis with Varimax rotation. Six components exceeded the eigenvalue-greater-than-one criterion. Their interpretability is assessed cautiously because this retention rule alone is not conclusive evidence of six distinct latent dimensions.

Validity and Reliability Test Result

Corrected item–total correlations and alpha-if-item-deleted statistics were examined to assess item functioning. All 275 responses were included in the reported analyses. As shown in Table 1, item quality varied; a strong total-score alpha should therefore not be interpreted as evidence that every item performed satisfactorily.

All 275 responses were included in the reported statistical analyses. The absence of excluded cases does not establish that each response was carefully considered or that the instrument was unambiguous; these aspects require evidence beyond missing-data counts.

The 25-item total score produced Cronbach's alpha of 0.892, indicating high internal consistency for this sample. This is evidence about the consistency of the total score, not proof of unidimensionality or of validity for all proposed component scores.

Corrected item–total correlations ranged from 0.158 to 0.657 (Table 1). Four items fell below the 0.30 screening criterion: Q1 (0.287), Q2 (0.158), Q24 (0.267), and Q25 (0.295). The alpha-if-deleted statistics for these items were 0.893, 0.896, 0.894, and 0.893, respectively, compared with the overall alpha of 0.892. Thus, contrary to the initial narrative, deleting Q2 or Q24 would marginally increase the reported alpha, although neither change alone provides a sufficient basis for deletion. All four items warrant further investigation of wording, content relevance, and distractor functioning before the instrument is used for consequential decisions.

Q10 (corrected item–total $r = 0.657$) and Q15 ($r = 0.651$) were among the strongest items in Table 1. Their alpha-if-deleted values were both 0.884, below the total-score alpha of 0.892. These statistics support retaining the items in the present version, subject to a broader review of their conceptual coverage.

Table 1
Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-total Correlation	Cronbach's Alpha if Item Deleted
Q1	17.9527	29.177	.287	.893
Q2	17.9927	29.766	.158	.896
Q3	18.1455	28.212	.441	.889
Q4	17.7745	28.876	.516	.887
Q5	17.9564	29.056	.310	.892
Q6	17.8036	28.348	.608	.885
Q7	17.9236	28.334	.480	.888
Q8	17.8655	28.190	.562	.886
Q9	17.8000	28.591	.549	.886
Q10	17.8145	28.093	.657	.884
Q11	17.8618	28.039	.602	.885
Q12	17.8473	28.466	.515	.887
Q13	17.7927	28.698	.532	.887
Q14	17.8509	28.069	.608	.885
Q15	17.8109	28.139	.651	.884
Q16	17.9709	28.379	.445	.889
Q17	17.8364	28.407	.543	.886
Q18	17.8073	28.521	.555	.886
Q19	17.9164	28.157	.524	.887
Q20	17.9018	28.103	.547	.886
Q21	17.9273	28.323	.480	.888
Q22	17.8400	28.069	.623	.885
Q23	17.8764	28.985	.366	.890
Q24	18.2182	29.135	.267	.894
Q25	18.1345	28.971	.295	.893

Exploratory Factor Analysis Test Result

Principal component analysis (PCA), as labelled in the available SPSS tables, was used to summarise the response covariation across the 25 assessment items. The six retained components offer a provisional descriptive grouping of proof-related reasoning items; they should not be interpreted as a demonstrated hierarchy of learners' cognitive abilities.

Sampling adequacy was high ($KMO = 0.906$), and Bartlett's test of sphericity was statistically significant ($\chi^2(300) = 2263.555$, $p < 0.001$). These results support proceeding with a dimension-reduction analysis but do not establish the number or conceptual validity of the components.

Reported communalities ranged from 0.445 to 0.725, indicating that the retained component solution accounted for different proportions of individual items' observed variance. These results should be interpreted in conjunction with the loading patterns, conceptual fit, and weaker item-total correlations rather than being treated as proof that every item has satisfactory construct validity.

Six components had initial eigenvalues greater than 1 (Table 2), collectively accounting for 56.628% of the total item variance. The first unrotated component explained 31.150%; after Varimax rotation, the rotated sums of squared loadings indicated a more distributed variance pattern (17.556%, 12.902%, 7.747%, 6.763%, 6.326%, and 5.333%). Because a scree plot is not reproduced, and no parallel analysis is reported, the six-component solution should be treated as exploratory.

Table 2
Total Variance Explained

Item	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.788	31.150	31.150	7.788	31.150	31.150	4.389	17.556	17.556
2	1.478	5.912	37.062	1.478	5.912	37.062	3.225	12.902	30.458
3	1.405	5.619	42.680	1.405	5.619	42.680	1.937	7.747	38.205
4	1.334	5.337	48.018	1.334	5.337	48.018	1.691	6.763	44.968
5	1.110	4.440	52.458	1.110	4.440	52.458	1.582	6.326	51.294
6	1.042	4.170	56.628	1.042	4.170	56.628	1.333	5.333	56.628
7	.964	3.857	60.485						
8	.874	3.497	63.982						
9	.854	3.415	67.397						
10	.774	3.096	70.493						
11	.731	2.924	73.417						
12	.675	2.700	76.118						
13	.635	2.540	78.657						
14	.594	2.377	81.034						
15	.564	2.255	83.289						
16	.552	2.210	85.499						
17	.527	2.108	87.608						
18	.468	1.873	89.481						
19	.461	1.844	91.325						
20	.429	1.716	93.042						
21	.393	1.573	94.615						
22	.374	1.497	96.111						
23	.371	1.482	97.594						
24	.311	1.244	98.838						
25	.291	1.162	100.000						

Extraction Method: Principal Component Analysis

The Varimax-rotated component matrix (Table 3) was interpreted using the largest loading for each item, while noting secondary loadings of at least 0.40. The following descriptive grouping is based on the actual coefficients in Table 3 and does not claim that the six components are independent psychological constructs:

- Component 1 (deductive and logical relationships): Q4, Q6, Q8, Q9, Q10, Q11, Q12, Q13, Q14, and Q15 had their largest loading on this component (approximately 0.467–0.746).
- Component 2 (linking premises through conditional reasoning): Q17, Q18, Q19, Q20, and Q22 had their largest loading on this component (approximately 0.520–0.723).
- Component 3 (identifying premises and conclusions): Q1, Q16, and Q21 had their largest loading on this component (approximately 0.692–0.711).
- Component 4 (recognising circular reasoning): Q24 and Q25 had their largest loading on this component (0.814 and 0.812, respectively).
- Component 5 (reasoning with particular propositions): Q5 and Q7 had their largest loading on this component (0.815 and 0.522). Q3 had no loading reaching the prespecified 0.40 threshold (its largest loading was 0.378), so its placement is unresolved.
- Component 6 (reasoning with universal propositions): Q2 and Q23 had their largest loading on this component (0.755 and 0.515, respectively).

Several items showed meaningful secondary loadings: Q12 loaded 0.482 on Component 1 and 0.411 on Component 6; Q15 loaded 0.579 on Component 1 and 0.401 on Component 2; and Q22 loaded 0.520 on Component 2 and 0.438 on Component 1. In contrast, Q18 primarily loaded on Component 2 (0.723), without a second loading of 0.40 or higher. The previously stated cross-loading of Q18 is therefore not supported by the displayed matrix. These patterns and the weak performance of Q3 require further validation.

Table 3
Rotated Component Matrix

	Component					
	1	2	3	4	5	6
Q1	.202	-.103	.711	-.039	.108	-.008
Q2	.120	-.027	-.137	.010	.136	.755
Q3	.197	.370	.150	.226	.378	-.230
Q4	.615	.076	.094	.115	.098	.207
Q5	.080	.158	.054	-.134	.815	.097
Q6	.644	.333	.109	.005	.169	-.105
Q7	.339	-.012	.210	.317	.522	.128
Q8	.493	.255	.058	.053	.341	.162
Q9	.746	.182	-.046	.030	.041	.069
Q10	.467	.350	.224	.127	.344	.103
Q11	.667	.100	.162	.113	.231	.068
Q12	.482	.101	.158	.156	.130	.411
Q13	.676	.215	.187	-.011	-.079	-.070
Q14	.586	.280	.219	.147	.027	.080
Q15	.579	.401	.189	-.001	.118	.106
Q16	.068	.314	.692	.081	.071	.139
Q17	.255	.596	.122	.225	-.047	.187
Q18	.238	.723	.009	.061	.201	-.021
Q19	.153	.716	.131	.044	.200	-.045
Q20	.355	.667	.053	-.076	-.036	.198
Q21	.241	.228	.700	.118	.067	-.111
Q22	.438	.520	.136	.147	.069	.065
Q23	.092	.366	.232	.054	-.057	.515
Q24	.077	.071	.135	.814	-.031	-.042
Q25	.105	.115	-.051	.812	.060	.114

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalisation.^a

^a Rotation converged in seven iterations.



Discussion

Validity and Reliability

The total-score Cronbach's alpha of 0.892 indicates high internal consistency in this sample. It does not demonstrate that all 25 items measure one latent trait or that the assessment directly measures the ability to write complete geometric proofs. Because the questions use selected-response formats, the scores are best interpreted as evidence of recognition and application of particular proof-related reasoning moves.

Table 1 suggests uneven item functioning. Four items (Q1, Q2, Q24, and Q25) had corrected item-total correlations below 0.30 and removing Q2 would raise alpha slightly from 0.892 to 0.896. These findings call for item-level review, not automatic deletion: low correlations may reflect item wording, response distributions, conceptual differences, or other factors that cannot be distinguished from the summary statistics alone (Finch, 2023). Importantly, these findings suggest that the instrument captures several aspects of students' reasoning processes, rather than measuring isolated procedural skills. This is consistent with theoretical perspectives on mathematics education that view proof construction as a complex cognitive activity involving multiple interconnected reasoning processes (Miyazaki et al., 2017; Stylianides et al., 2024).

The Factor Structure

The rotated PCA solution suggests six descriptive clusters of responses related to reasoning in geometric proof tasks. These components are analytical summaries, not collaborative or necessarily separate cognitive processes. Although their proposed labels are broadly compatible with scholarship on deductive proof and comprehension (Miyazaki et al., 2017; Yang & Lin, 2008), the labels remain interpretive and need verification with independent samples and more appropriate latent-variable analyses.

The first component groups items concerned with drawing deductive conclusions and recognising logical relationships between geometric statements. It may be relevant when teacher educators ask prospective teachers to analyse whether a conclusion follows from the premises supplied. However, component loadings alone cannot show that high-scoring participants could independently construct coherent written proofs.

The second component concerns linking premises through conditional reasoning, including hypothetical syllogism: if P implies Q and Q implies R , then P implies R . Identifying this chain of implications is important when analysing whether intermediate statements logically connect a geometric argument. For mathematics teacher education, these tasks may support discussions of missing justifications and invalid transitions between proof steps.

The third component concerns identifying relevant premises and formulating appropriate conclusions. Anwar et al. (2021) and Yang and Lin (2008) discussed difficulties associated with recognising the logical organisation and meaning of proof statements. Prospective teachers could use this distinction when evaluating whether a learner has selected relevant givens and stated a conclusion justified by the available information.

The fourth component comprised items targeting the recognition of circular arguments, in which the intended conclusion is assumed or used as its own justification. This is a useful topic for preservice teacher discussion, although the two-item component and the low item-total correlation for Q24 require caution before interpreting it as a reliable subscale.

The fifth component is provisionally interpreted as reasoning with particular propositions-statements concerning specific geometric objects or configurations. These statements may form intermediate steps in a deductive argument (Miyazaki et al., 2017). Q3 did not reach the prespecified loading threshold, so the content coverage and distinctiveness of this component require further study.

The sixth component is provisionally concerned with recognising and using universal propositions (Soames, 1995) in particular geometric cases. In analysing such reasoning, preservice teachers may need to distinguish a general theorem from its application to a specific diagram. Because this component contains only two items, including Q2 with a low item–total correlation, it should not yet be treated as a validated independent score.

Taken together, the six components suggest several potentially useful areas for examining preservice teachers' attention to the logical organisation of geometric arguments. The component solution is descriptive rather than developmental: it does not establish a necessary order in which proof skills are acquired, nor does it independently validate six teachable or assessable subscales.

The reported response patterns are broadly consistent with theoretical discussions of geometric proof reasoning and proof comprehension (Miyazaki et al., 2017; Yang & Lin, 2008). Nevertheless, an exploratory component analysis alone cannot validate a fully specified multidimensional model. The present study contributes preliminary instrument evidence and identifies candidate domains for item refinement and for subsequent investigation in mathematics teacher education.

Implications for Mathematics Education

The findings have implications not only for the teaching and assessment of geometric proof, but also for mathematics teacher education. The identification of multiple components underlying students' proof construction—including logical relationships, the use of logic, and universal claims—provides mathematics teacher educators with specific aspects of proof and proving around which preservice and practising teachers' knowledge for teaching can be developed. Research in mathematics teacher education has demonstrated that teaching proof requires more than teachers' own knowledge of formal mathematical proof; teachers also need knowledge of how students understand, construct, and struggle with proofs and how instructional activities can support the development of this understanding (Buchbinder & McCrone, 2020; Leung & Lee, 2017). Thus, the components identified in this study can be considered not simply as features to be measured in students' work, but as potential focus points for developing teachers' capacity to interpret and respond to students' reasoning.

From this perspective, mathematics teacher education should provide preservice teachers with opportunities to examine the constituent processes involved in proof construction and to connect these processes with instructional decision-making. Rather than treating proof as a finished formal product, teacher education can foreground the processes through which learners identify relevant propositions and premises, establish logical relationships between statements, evaluate the adequacy of arguments, and determine whether a claim has been established universally. Such an approach is consistent with research showing that preservice teachers can experience difficulties in recognising and teaching the logical aspects of proof, including quantification, conditional statements, direct and indirect reasoning, and argument evaluation (Buchbinder & McCrone, 2020). Recent research similarly demonstrates the value of reasoning-and-proving tasks that explicitly connect preservice teachers' experiences of mathematical reasoning with the work of teaching and that support teacher development through cycles of exploration, argumentation, and proof (Lee, 2026).

The present 25-item selected-response assessment may serve as an exploratory diagnostic resource for mathematics teacher educators. For example, teacher educators could ask prospective teachers to justify each selected answer, examine why distractors are invalid, and compare students' explanations of circular arguments, conditional reasoning, and universal claims. Such activities would extend beyond the scores themselves by making the underlying mathematical reasoning visible. This proposed use is pedagogically plausible but was not empirically evaluated in the present validation study.

The findings also suggest that mathematics teacher education should move beyond an exclusive emphasis on teaching formal proof procedures. Research on preservice and novice teachers indicates that teachers may have difficulty using forms of proof that connect formal mathematical structures with meaningful geometric reasoning (Leung & Lee, 2017). Consequently, teacher education experiences could engage preservice teachers in analysing and comparing different forms of justification, including



informal and pre-formal arguments, empirical examples, deductive arguments, and formal proofs. This would help preservice teachers develop a more nuanced understanding of what constitutes mathematical justification and how different forms of reasoning can contribute to students' progression towards formal proof.

More broadly, the findings support positioning proof as a process of reasoning and justification rather than as a mechanical sequence of steps to be reproduced. Mathematics teacher education can contribute to this shift by giving preservice teachers opportunities to construct proofs themselves, analyse the reasoning embedded in existing proofs, anticipate students' difficulties, and design tasks that make logical relationships visible. Such experiences can be embedded within school mathematics topics, including geometry, rather than treating proof as a specialised activity confined to a particular unit of the curriculum. Buchbinder and McCrone (2020), for example, demonstrated the potential of teacher education that combines attention to the logical aspects of proof with lesson planning and classroom enactment, enabling preservice teachers to develop mathematical knowledge for teaching proof through practice.

The contribution to mathematics teacher education is therefore a provisional map of reasoning features that teacher educators could use when designing tasks about students' geometric arguments. Before adopting component-specific scores for placement, evaluation, or high-stakes decisions, further testing is required, including independent-sample structural validation, additional items for sparsely represented components, and studies connecting assessment results to teachers' instructional decisions. The present data do not demonstrate improvements in teachers' practice or in secondary students' proof learning.

Conclusion

This study evaluated a 25-item selected-response instrument intended to assess aspects of geometric proof-related reasoning among 275 Indonesian preservice mathematics teachers. Its total score showed high internal consistency (Cronbach's $\alpha = 0.892$), and the reported principal component analysis produced a provisional six-component solution accounting for 56.628% of item variance. These findings are promising but are not sufficient to establish six validated latent dimensions or the ability to construct complete written proofs.

For mathematics teacher education, the instrument highlights reasoning features that prospective teachers may need to recognise in learners' arguments, including justified deductions, conditional links, relevant premises, circular reasoning, and particular versus universal claims. Teacher educators may use the items as prompts for analysing students' reasoning and planning instructional responses; such pedagogical applications should be tested rather than assumed.

Important limitations include four items with corrected item-total correlations below 0.30, an item without a loading meeting the stated 0.40 threshold, cross-loadings, two-item components, and the use of PCA rather than a confirmed latent-factor model in the reported tables. The multiple-choice response format further limits claims about independent written-proof construction. Subsequent work should revise and pilot weaker items, verify the extraction method, assess dimensionality with appropriate techniques, and evaluate evidence for practical teacher-education uses.

The instrument therefore provides preliminary evidence for examining aspects of proof-related reasoning, rather than a fully validated multidimensional measure ready for all instructional purposes. Its potential value lies in connecting assessment evidence to teacher educators' explicit discussion of how learners justify geometric statements and how teachers might respond to incomplete or circular reasoning. This link remains a direction for future research, not an outcome established by the present data.

Corresponding author

Lathiful Anwar
 Department of Mathematics,
 Universitas Negeri Malang
 Jalan Semarang No 5 Malang 65145 East Java Indonesia
 Lathiful.anwar.fmipa@um.ac.id

Ethical approval

Ethical approval for this study was obtained from the Research Ethics Committee of Universitas Negeri Malang (Ref. No. 30.9.1/UN32.14/PB/2024). All participants agreed voluntarily to participate in the study and provided informed consent prior to data collection. Participants were informed about the purpose of the study, the voluntary nature of their participation, and their right to withdraw at any time. All data were treated confidentially, and anonymised prior to analysis to ensure participants' privacy.

Funding

This investigation was sponsored by the Fundamental research scheme of the Direktorat Riset, Teknologi, dan Pengabdian Kepada Masyarakat (DRTPM) under contract No. 11.6.88 /UN32.14.1/LT/2024.

Competing interests

The authors declare there are no competing interests.

References

- Anwar, L., Goedhart, M. J., & Mali, A. (2023). Learning trajectory of geometry proof construction: Studying the emerging understanding of the structure of Euclidean proof. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(5), Article em2266. <https://doi.org/10.29333/ejmste/13160>
- Anwar, L., Mali, A., & Goedhart, M. J. (2021). The effect of proof format on reading comprehension of geometry proof: The case of Indonesian prospective mathematics teachers. *Eurasia Journal of Mathematics, Science and Technology Education*, 17(4), Article em1952. <https://doi.org/10.29333/EJMSTE/10782>
- Buchbinder, O., & McCrone, S. (2020). Preservice teachers learning to teach proof through classroom implementation: Successes and challenges. *The Journal of Mathematical Behavior*, 58, Article 100779. <https://doi.org/10.1016/j.jmathb.2020.100779>
- Cirillo, M., & Hummer, J. (2021). Competencies and behaviors observed when students solve geometry proof problems: An interview study with smartpen technology. *ZDM—Mathematics Education*, 53(4), 861–875. <https://doi.org/10.1007/s11858-021-01221-w>
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Taylor & Francis.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
- Field, A. P. (2017). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE Publications.
- Finch, W. H. (2023). A comparison of methods for determining the number of factors to retain with exploratory factor analysis of dichotomous data. *Psych*, 5(3), 1004–1018. <https://doi.org/10.3390/psych5030067>



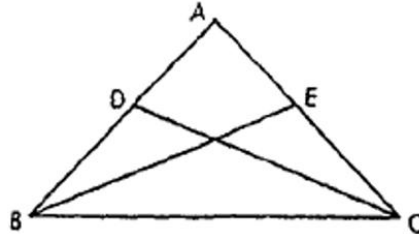
- Fujita, T., & Jones, K. (2014). Reasoning-and-proving in geometry in school mathematics textbooks in Japan. *International Journal of Educational Research*, 64, 81–91. <https://doi.org/10.1016/j.ijer.2013.09.014>
- Fujita, T., Jones, K., & Miyazaki, M. (2018). Learners' use of domain-specific computer-based feedback to overcome logical circularity in deductive proving in geometry. *ZDM–Mathematics Education*, 50(4), 699–713. <https://doi.org/10.1007/s11858-018-0950-4>
- Hair, J., Anderson, R., Babin, B., & Black, W. (2013). *Multivariate data analysis: Pearson new international edition* (7th ed.). Pearson Deutschland. <https://elibrary.pearson.de/book/99.150005/9781292035116>
- Hohol, M., & Miłkowski, M. (2019). Cognitive artifacts for geometric reasoning. *Foundations of Science*, 24(4), 657–680. <https://doi.org/10.1007/s10699-019-09603-w>
- Jablonski, S., & Ludwig, M. (2023). Teaching and learning of geometry: A literature review on current developments in theory and practice. *Education Sciences*, 13(7), Article 682. <https://doi.org/10.3390/educsci13070682>
- Jones, K., Miyazaki, M., & Fujita, T. (2015). Aspects of scaffolding in a web-based learning system for congruency-based proofs in geometry. In N. Amado & S. Carreira (Eds.), *Proceedings of the 12th International Conference on Technology in Mathematics Teaching (ICTMT12)*, Portugal (pp. 561–563). Universidade do Algarve.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. <https://doi.org/10.1007/BF02291575>
- Komatsu, K., & Jones, K. (2019). Task design principles for heuristic refutation in dynamic geometry environments. *International Journal of Science and Mathematics Education*, 17(4), 801–824. <https://doi.org/10.1007/s10763-018-9892-0>
- Lee, G. C.-Y. (2026). Using reasoning-and-proving cycles to foster pre-service teacher development in argumentation and proof: An illustration of practice. *Mathematics Teacher Education and Development*, 28(1), IoP Article 1.
- Leung, K. C. I., & Lee, C. Y. (2017). Preservice and novice teachers' knowledge on preformal proofs: Triangle postulate as an example. *Mathematics Teacher Education and Development*, 19(2), 51–80.
- Makridis, O. (2022). Concepts of deductive reasoning. In O. Makridis (Ed.), *Symbolic logic* (pp. 57–80). Springer International Publishing. https://doi.org/10.1007/978-3-030-67396-3_2
- Miyazaki, M., & Fujita, T. (2015). Proving as an explorative activity in mathematics education: New trends in Japanese research into proof. In B. Sriraman, J. Cai, K.-H. Lee, L. Fan, Y. Shimizu, C. S. Lim, & K. Subramaniam (Eds.), *The first sourcebook on Asian research in mathematics education: China, Korea, Singapore, Japan, Malaysia and India* (pp. 1375–1407). Information Age Publishing.
- Miyazaki, M., Fujita, T., & Jones, K. (2014). Functions of open flow-chart proving in introductory lessons of formal proof. In P. Liljedahl, S. Oesterle, C. Nicol, & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (Vol. 4, pp. 225–232). PME.
- Miyazaki, M., Fujita, T., & Jones, K. (2015). Flow-chart proofs with open problems as scaffolds for learning about geometrical proofs. *ZDM–Mathematics Education*, 47(7), 1211–1224. <https://doi.org/10.1007/s11858-015-0712-5>
- Miyazaki, M., Fujita, T., & Jones, K. (2017). Students' understanding of the structure of deductive proof. *Educational Studies in Mathematics*, 94(2), 223–239. <https://doi.org/10.1007/s10649-016-9720-9>
- Miyazaki, M., Nagata, J., Chino, K., Sasa, H., Fujita, T., Komatsu, K., & Shimizu, S. (2019). Curriculum development for explorative proving in lower secondary school Geometry: Focusing on the levels of planning and constructing a proof. *Frontiers in Education*, 4. <https://doi.org/10.3389/feduc.2019.00031>
- Ramírez-Uclés, R., & Ruiz-Hidalgo, J. F. (2022). Reasoning, representing, and generalizing in geometric proof problems among 8th grade talented students. *Mathematics*, 10(5), Article 789. <https://doi.org/10.3390/math10050789>
- Shrestha, N. (2021). Factor analysis as a tool for survey analysis. *American Journal of Applied Mathematics and Statistics*, 9(1), 4–11. <https://doi.org/10.12691/ajams-9-1-2>
- Soames, S. (1995). Beyond singular propositions? *Canadian Journal of Philosophy*, 25(4), 515–549. <https://doi.org/10.1080/00455091.1995.10717425>
- Stein, X. V., Tsarava, K., & Goecke, B. (2026). Development and validation of a preformal test for mathematical proof competence. *Thinking Skills and Creativity*, 61, Article 102139. <https://doi.org/10.1016/j.tsc.2026.102139>
- Stylianides, G. J., Stylianides, A. J., & Moutsios-Rentzos, A. (2024). Proof and proving in school and university mathematics education research: A systematic review. *ZDM–Mathematics Education*, 56(1), 47–59. <https://doi.org/10.1007/s11858-023-01518-y>
- Wu, A. D. (2020). Varimax loadings. In F. Maggino (Ed.), *Encyclopedia of quality of life and well-being research*. Springer International Publishing. https://doi.org/10.1007/978-3-319-69909-7_3150-2
- Yang, K. L., & Lin, F. L. (2008). A model of reading comprehension of geometry proof. *Educational Studies in Mathematics*, 67(1), 59–76. <https://doi.org/10.1007/s10649-007-9080-6>



Appendix: Test Items

Read and understand the following statement:

If in triangle ABC, as shown in the figure, side DB is congruent to side EC ($DB \cong EC$) and side DC is congruent to side EB ($DC \cong EB$), then angle $\angle BDC$ is congruent to angle $\angle CEB$ ($\angle BDC \cong \angle CEB$).



Answer the following multiple-choice questions:

Note: You can create a proof of the truth of the statement above before answering the questions below.

Multiple-Choice Questions:

1. To prove that $\angle BDC \cong \angle CEB$, which specific statement needs to be proven first?
 - A) $\triangle BDC \cong \triangle CEB$
 - B) $DB \parallel EC$
 - C) D and E are the midpoints of BC
 - D) $\triangle ABD \cong \triangle AEC$

2. If $DB \cong EC$ and $DC \cong EB$, which general statement supports the conclusion that $\triangle BDC \cong \triangle CEB$?
 - A) If two triangles have three equal sides, then both triangles are congruent.
 - B) If two triangles have two equal sides and corresponding angles of equal size, then both triangles are congruent.
 - C) If two angles in a triangle are equal, then the triangle is similar.
 - D) If a triangle has a right angle, then the angles are complementary.

3. Which statement is necessary to show that $\triangle BDC \cong \triangle CEB$?
 - A) Angles $\angle B$ and $\angle C$ are congruent.
 - B) Angles $\angle BDC \cong \angle CEB$
 - C) Triangles $\triangle BDC$ and $\triangle CEB$ are isosceles.
 - D) Sides $DB \cong EC$ and $DC \cong EB$

4. Why do we need to prove that $\triangle BDC \cong \triangle CEB$ to show that $\angle BDC \cong \angle CEB$?
 - A) Because triangle congruence ensures that corresponding angles are equal.
 - B) Because opposite angles in a triangle are always equal.
 - C) Because all angles in an equilateral triangle are equal.
 - D) Because the internal angles of a triangle always sum to 180 degrees.

5. To show that $\triangle BDC \cong \triangle CEB$, which general statement should we use?
- A) A triangle with two equal sides is isosceles.
 - B) If two triangles are congruent, all corresponding sides are equal in length.
 - C) The angles opposite equal sides are equal in size.
 - D) Any triangle with three equal sides is equilateral.
6. If we want to prove that $\triangle BDC \cong \triangle CEB$, which specific statement needs to be shown first?
- A) Sides $BD \cong EC$ and $DC \cong EB$.
 - B) Angle $\angle BDC$ is a right angle.
 - C) Line BC is bisected by line AD .
 - D) Angle $\angle BDC \cong \angle BAC$
7. Which statement is most relevant to use in proving that angle $\angle BDC \cong \angle CEB$?
- A) If two triangles are congruent, their corresponding angles are equal.
 - B) Every isosceles triangle has equal angles.
 - C) If opposite angles in a triangle are equal, the triangle is similar.
 - D) If two angles in a triangle are equal, the opposite sides are also equal.
8. If $DB \cong EC$ and $DC \cong EB$, what can we conclude about triangles $\triangle BDC$ and $\triangle CEB$?
- A) Both triangles are similar.
 - B) Both triangles are congruent.
 - C) Both triangles are equilateral.
 - D) Both triangles are scalene.
9. What conclusion can we draw if it is known that $\triangle BDC \cong \triangle CEB$?
- A) Side BC is the same length as side AD
 - B) Line AD is perpendicular to BC
 - C) Angle $\angle BDC \cong \angle CEB$
 - D) Side DB is longer than side EC
10. Why is it necessary to show that sides $BD \cong EC$ and $DC \cong EB$ to prove that $\triangle BDC \cong \triangle CEB$?
- A) Because the corresponding sides in congruent triangles are equal.
 - B) Because angles in a triangle must always be equal.
 - C) Because isosceles triangles always have two equal sides.
 - D) Because every triangle has three equal sides.
11. If we know that $\triangle BDC \cong \triangle CEB$, what statement must we make about the corresponding angles?
- A) Angles $\angle BDC$ and $\angle CEB$ are equal.
 - B) Angles $\angle BDC$ and $\angle CEB$ are not equal.
 - C) Angle $\angle BDC$ is obtuse.
 - D) Angles $\angle BDC$ and $\angle CEB$ are right angles.
12. Why is it important to show that $\triangle BDC \cong \triangle CEB$ before concluding that $\angle BDC \cong \angle CEB$?
- A) Because triangle congruence ensures that corresponding angles are equal.
 - B) Because all angles in a triangle must always be equal.
 - C) Because congruent triangles must have equal sides.
 - D) Because equal angles only occur in scalene triangles.



13. What can be concluded if $DB \cong EC$ and $DC \cong EB$?
- Triangles $\triangle BDC$ and $\triangle CEB$ are congruent.
 - Angles $\angle BDC$ and $\angle CEB$ are not equal.
 - Line AD bisects the triangle into two equal parts.
 - Triangles $\triangle ABD$ and $\triangle AEC$ are congruent.
14. To prove that angle $\angle BDC \cong \angle CEB$, which specific statement must be shown first?
- Triangles $\triangle BDC \cong \triangle CEB$
 - Line BC is straight
 - Side $AB \cong AC$
 - Angle $\angle BAC \cong \angle BCA$
15. What is the relationship between the sides and angles of the triangles if $\triangle BDC \cong \triangle CEB$?
- Corresponding sides and angles are equal in length and size.
 - Only corresponding sides are equal in length.
 - Only corresponding angles are equal in size.
 - All sides and angles differ in both triangles.
16. What is the most appropriate first step to prove that $\angle BDC \cong \angle CEB$?
- Show that angles $\angle BDA \cong \angle CEA$.
 - Show that $DB \parallel EC$.
 - Show that $\triangle BDC \cong \triangle CEB$.
 - Show that D and E are midpoints of BC.
17. If it is known that $DB \cong EC$ and $DC \cong EB$, which specific statement is most relevant to use in proving that $\triangle BDC \cong \triangle CEB$?
- Side $AB \cong AC$.
 - Angle $\angle BDC$ is obtuse.
 - Line AD is perpendicular to BC.
 - Triangles $\triangle BDC$ and $\triangle CEB$ have corresponding sides of equal length.
18. Which statement indicates that angle $\angle BDC \cong \angle CEB$ based on the given information?
- If D and E are the midpoints of BC, then $\angle BDC \cong \angle CEB$
 - If $DB \cong EC$ and $DC \cong EB$, then triangle $\triangle BDC \cong \triangle CEB$
 - If line AD is perpendicular to BC, then angle $\angle BDC \cong \angle CEB$
 - If $AB \cong AC$, then $\triangle ABD \cong \triangle AEC$
19. Given:
- Statement 1: If $DB \cong EC$ and $DC \cong EB$, then triangle $\triangle BDC \cong \triangle CEB$.
- Statement 2: If $\triangle BDC \cong \triangle CEB$, then the corresponding angles are equal.
- What conclusion can we draw from these two statements?
- If $DB \cong EC$ and $DC \cong EB$, then angle $\angle BAC \cong \angle ABC$
 - If $DB \cong EC$ and $DC \cong EB$, then line AD is perpendicular to BC
 - If $DB \cong EC$ and $DC \cong EB$, then $\angle BDC \cong \angle CEB$
 - If $DB \cong EC$ and $DC \cong EB$, then triangle $\triangle ABD \cong \triangle AEC$



20. Given:

Statement 1: If $\triangle BDC \cong \triangle CEB$, then all corresponding sides and angles are equal.

Statement 2: If $DB \cong EC$ and $DC \cong EB$, then $\triangle BDC \cong \triangle CEB$.

What conclusion can we draw from combining these two statements?

- A) If $DB \cong EC$ and $DC \cong EB$, then angle $\angle BAC \cong \angle ABC$.
- B) If $DB \cong EC$ and $DC \cong EB$, then the triangle is scalene.
- C) If $DB \cong EC$ and $DC \cong EB$, then line AD bisects BC into two equal parts.
- D) If $DB \cong EC$ and $DC \cong EB$, then all corresponding sides and angles in the triangle are equal.

21. To prove that $\angle BDC \cong \angle CEB$, which single statement needs to be proven first?

- A) $\triangle BDC \cong \triangle CEB$
- B) $DB \parallel EC$
- C) $\angle BDA \cong \angle CEA$
- D) D and E are the midpoints of BC

22. Which single statement is needed to prove that $\triangle BDC \cong \triangle CEB$?

- A) Line AD is the altitude of the triangle
- B) Angle $\angle BDA \cong \angle CEA$
- C) $DB \cong EC$ and $DC \cong EB$
- D) Angle $\angle BDA$ is greater than angle $\angle CEA$

23. If we know that $DB \cong EC$ and $DC \cong EB$, which single statement needs to be proven to support that $\triangle BDC \cong \triangle CEB$?

- A) Angles $\angle B$ and $\angle C$ are congruent
- B) Line AD is perpendicular to BC
- C) Triangles $\triangle BDC$ and $\triangle CEB$ have corresponding sides of equal length
- D) Angle $\angle BAC \cong \angle ABC$

24. What does the following incorrect proof state: "Since $\angle BDC \cong \angle CEB$, then $DB = EC$ and $DC = EB$, so $\angle BDC \cong \angle CEB$ "?

- A) This proof is correct because it uses triangle properties.
- B) This proof assumes the result that needs to be proven.
- C) This proof is wrong because it does not use additional information.
- D) This proof is wrong because it does not use the corresponding angle theorem.

25. What logical error exists in the following proof: "Since $\triangle BDC \cong \triangle CEB$, then $\angle BDC \cong \angle CEB$, and because these angles are congruent, $\triangle BDC \cong \triangle CEB$ "?

- A) This proof is correct because it uses the properties of congruent triangles.
- B) This proof uses the result that needs to be proven as the basis for the proof.
- C) This proof is wrong because angles in a triangle are not always equal.
- D) This proof is wrong because it uses incorrect information.

