



Standardized Assessment of TKA Equivalent for Junior High School Students

Nur Roudlotul Jannah^{1,*}, Cholis Sa'dijah², Tjang Daniel Chandra³

^{1, 2, 3} Universitas Negeri Malang, Malang, Indonesia

*Corresponding Authors: nur.roudlotul.2303118@students.um.ac.id

Submitted: 25-04-2026

Revised: 28-04-2026

Accepted: 29-04-2026

Published: 05-06-2026

ABSTRACT

This research aims to develop a mathematical assessment instrument equivalent to the Academic Ability Test (TKA) for junior high school students who have good measurement quality. This research uses a Research and Development (R&D) approach by adapting the ADDIE model which includes the stages of analysis, design, development, implementation, and evaluation. The instrument developed consists of three question items that are arranged based on the TKA preparation grid with emphasis on mathematical reasoning skills, problem solving, and critical thinking. The research subjects consisted of expert validators including experts in mathematics and educational evaluation experts, as well as 80 junior high school students as respondents to the instrument trial. The data collection technique is carried out through expert validation sheets and tests of student work. Data analysis was carried out descriptively to assess the feasibility of the instrument's contents, while empirical validity was analyzed using corrected item-total correlation and reliability was analyzed using Cronbach's Alpha coefficient with the help of IBM SPSS Statistics software. The results showed that all question items had a corrected item-total correlation value that was greater than the r -value of the table (0.220) so that it was declared valid. In addition, the results of the reliability test showed a Cronbach's Alpha value of 0.953 which was in the very high category. These findings show that the assessment instruments developed have a very good level of validity and reliability so that they are suitable for use as an alternative assessment to measure the academic ability of junior high school students that are in line with TKA standards.

Keywords: academic ability test; instrument development; mathematics assessment

INTRODUCTION

21st century education requires students to possess higher-order thinking skills, particularly problem-solving and critical thinking, as essential competencies to face the complexities of modern life (Jannah et al., 2026). In the context of junior high school education, the development of these abilities serves as a fundamental basis before students advance to higher levels of education. Mathematics plays a strategic role as a medium to train logical reasoning, analytical skills, and systematic problem-solving abilities (Rahman et al., 2021; Utami et al., 2025). Therefore, mathematics learning should be supported by an evaluation system that is capable of comprehensively describing students' cognitive processes and thinking skills. Learning evaluation not only functions to measure the level of material mastery, but also to provide meaningful information about how students think in solving problems (Phafiandita et al., 2022). Thus, appropriate assessment practices are essential to accurately measure the development of students' cognitive abilities.

In the context of educational evaluation, assessment is defined as a systematic process of collecting, analyzing, and interpreting information about students' abilities as a

basis for decision-making in learning (Adriantoni et al., 2025; Sa'dijah et al., 2025). A high-quality assessment must be able to measure the intended constructs accurately and consistently (Munaroh, 2024). Therefore, the development of assessment instruments must consider measurement quality, particularly in terms of validity and reliability. Validity refers to the extent to which an instrument measures what it is intended to measure, while reliability indicates the consistency of measurement results when the instrument is administered under similar conditions (Saputri & Larasati, 2023; Widiana, 2023). Instruments that lack validity or reliability may produce inaccurate data, which can lead to incorrect interpretations of students' abilities (Munaroh, 2024). Consequently, the development of assessment instruments must involve systematic validation and empirical testing to ensure their feasibility and accuracy in evaluating learning outcomes.

In line with the need for high-quality assessment, the Academic Ability Test (Tes Kemampuan Akademik/TKA) has been introduced as an assessment model that emphasizes reasoning skills. TKA is designed to measure students' academic abilities not only in terms of content mastery but also in understanding, analyzing, and solving problems (Kemendikdasmen, 2025). The TKA framework highlights reasoning ability, problem-solving, and numeracy literacy as key indicators of academic competence. The standards for TKA development issued by the Ministry of Primary and Secondary Education provide a structured and standardized reference for designing assessment instruments (Kemendikdasmen, 2025). These principles allow adaptation across different levels of education, including junior high school. Therefore, integrating TKA principles into assessment practices is expected to strengthen the measurement of students' academic abilities from an early stage.

However, the implementation of mathematics assessment at the junior high school level still faces several challenges. Existing assessment instruments are generally oriented toward procedural knowledge and memorization of concepts (Syafi et al., 2025). The items presented have not optimally measured complex reasoning skills or students' ability to transfer knowledge to new situations. As a result, the information obtained from assessment results does not fully reflect students' cognitive abilities comprehensively. Teachers require assessment instruments that are not only valid and reliable but also practical and aligned with national standards (Hidayat et al., 2026; Saputra & Akmal, 2025). In addition, the limited availability of instruments that are aligned with standardized frameworks such as TKA at the junior high school level indicates the need for more representative and high-quality assessment tools.

These conditions reveal a significant research gap in the field of mathematics assessment development. Most previous studies tend to focus on improving learning outcomes or examining the effectiveness of instructional models (Aqidatul Aisyah et al., n.d.; Paat et al., 2024; Ramadhani & Sukenti, 2023). Meanwhile, studies that specifically emphasize the development and empirical testing of assessment instrument quality, particularly those aligned with academic ability standards such as TKA, remain limited. Furthermore, research that integrates content validation, empirical validity testing, and reliability analysis within a single framework of TKA-equivalent assessment development for junior high school students has not been widely conducted. This indicates a clear need

for research that focuses on developing standardized, valid, and reliable assessment instruments that are capable of measuring higher-order thinking skills in mathematics.

Based on this gap, this study offers a novel contribution by developing a mathematics assessment instrument equivalent to TKA specifically for junior high school students and testing its quality through a comprehensive validation process. The instrument is designed based on the TKA framework, emphasizing the measurement of mathematical reasoning, problem solving, and critical thinking skills. The items are constructed to represent the indicators outlined in the TKA standards, enabling a more comprehensive assessment of students' academic abilities. The development process involves systematic stages, including item construction, expert validation, and empirical testing involving students. This approach ensures that the resulting instrument meets the criteria of a high-quality assessment tool and is suitable for use in mathematics learning evaluation.

Based on the description above, the research problems in this study focus on examining the quality of the developed assessment instrument. Specifically, this study seeks to investigate the content validity of the TKA-equivalent mathematics assessment instrument based on expert judgment, to analyze the empirical validity of the instrument based on students' responses, and to determine the reliability of the instrument in measuring students' academic abilities.

Accordingly, this study aims to develop a TKA-equivalent mathematics assessment instrument for junior high school students and to evaluate its quality in terms of content validity, empirical validity, and reliability. The results of this study are expected to demonstrate that the developed instrument has strong measurement quality and can be used as an alternative evaluation tool in mathematics learning. Ultimately, this research is expected to contribute to the advancement of educational assessment studies and support more effective evaluation practices that are oriented toward higher-order thinking skills.

RESEARCH METHODS

This research is a research and development (R&D) that aims to produce a mathematical assessment instrument equivalent to the Academic Ability Test (TKA) for junior high school students who have good measurement quality. The development design adapts the ADDIE model which consists of five stages, namely analysis, design, development, implementation, and evaluation with adjustments to the development of educational assessment instruments. The development stages in this study include Analysis, at this stage a study of the needs for instrument development is carried out through an examination of the TKA preparation grid published by the Ministry of Primary and Secondary Education and mathematics competency analysis at the junior high school level. This stage aims to identify relevant indicators of ability to be measured, in particular mathematical reasoning, problem-solving, and critical thinking skills. Next is Instrument Design, this stage includes the preparation of assessment grids based on the ability indicators contained in the TKA preparation standards. Furthermore, the design of question items is carried out that is adjusted to the ability indicators to be measured, as well as the determination of the form of the question and scoring guidelines. In the Instrument Development stage, three items of mathematical assessment questions are prepared in accordance with the indicators in the grid that has been designed. The question items that

have been developed are then validated by experts in mathematics and educational evaluation experts to assess the suitability of the content, construction, and linguistic aspects of the instrument. The results of expert validation are used as a basis for revising instruments before testing students. Furthermore, the Implementation stage (Instrument Trial) was carried out by testing instruments to junior high school students to obtain student response data to the developed question items. The data from the students' work is used as a basis for empirically analyzing the quality of the question items. The last stage is Evaluation by analyzing the quality of the instrument based on expert validation results and empirical test results. This evaluation aims to determine the feasibility of the assessment instrument developed so that it can be used as a measuring tool for students' academic ability equivalent to TKA.

The research subjects consisted of expert validators including experts in mathematics and educational evaluation experts, as well as 80 junior high school students as respondents to the instrument trial. The data collection technique is carried out through expert validation sheets and tests of student work results on the developed instruments. The data from expert validation results were analyzed descriptively to determine the feasibility level of the instrument's contents. Furthermore, the empirical validity of the question items was analyzed using the correlation between the item score and the total score to determine the degree of relevance of each question item to the overall instrument. The instrument reliability test was carried out using the Alpha Cronbach coefficient to determine the level of internal consistency of the developed assessment instrument. The entire statistical analysis was done with the help of IBM SPSS Statistics software. The results of the validity and reliability analysis are used as a basis to determine the feasibility of assessment instruments equivalent to TKA for junior high school students.

RESULTS AND DISCUSSION

The development of the TKA-equivalent mathematics assessment instrument began with a validation process conducted by material experts. This stage aimed to evaluate the feasibility of the instrument in terms of content, construction, and language aspects. The results of expert judgment indicated that the instrument was generally appropriate for use, although several revisions were required. These revisions primarily focused on improving the clarity of item wording, refining the contextual representation of problems, and aligning the indicators more closely with the TKA framework. This validation process plays a crucial role in ensuring that the instrument meets fundamental quality standards before being implemented in empirical testing, as expert validation is essential for establishing content validity in assessment development.

Following the validation stage, empirical testing was conducted involving 80 junior high school students to examine the validity of the instrument. The analysis used corrected item-total correlation to determine the extent to which each item contributes to the overall construct being measured. As presented in Table 1, all items demonstrate corrected item-total correlation values exceeding the critical r -value (0.220), indicating that each item has a strong relationship with the total score.

Table 1. Results of the Empirical Validity Test

	Corrected Item- Total Correlation	Cronbach's Alpha if Item Deleted
Item 1	0.870	0.955
Item 2	0.943	0.900
Item 3	0.891	0.939

These findings indicate that all items are valid and effectively represent the intended construct of mathematical ability. The strong correlation values suggest that each item consistently contributes to measuring the same underlying construct, namely students' reasoning and problem-solving abilities. This result can be attributed to the systematic development process, which involved alignment with the TKA framework and refinement through expert validation. Such a process ensures that the items are not only relevant but also capable of capturing higher-order thinking skills. This finding is consistent with previous studies which state that assessment instruments developed through structured validation and empirical testing tend to demonstrate strong validity and accurately represent the intended constructs (Saputri & Larasati, 2023; Ulfah, 2025; Zayrin et al., 2025).

In addition to validity, the reliability of the instrument was analyzed using Cronbach's Alpha to determine its internal consistency. The results, as shown in Table 2, indicate that the instrument has a very high reliability coefficient.

Table 2. Reliability Test Results

N of Item	Cronbach's Alpha
3	0.953

The high reliability coefficient demonstrates that the instrument produces consistent measurement results and that all items function cohesively in assessing the same construct. This strong internal consistency can be explained by the alignment of items with clearly defined indicators derived from the TKA framework, as well as the careful construction of items during the development process. Instruments with high reliability are considered stable and dependable when used across similar groups of respondents. This finding supports previous research indicating that well-constructed assessment instruments tend to yield consistent and reliable measurements of students' abilities (Muharromah et al., 2025; Saputri & Larasati, 2023; Zana et al., 2025).

Beyond statistical evidence, the characteristics of the developed items indicate that the instrument has successfully incorporated higher-order thinking skills. The items require students not only to apply mathematical procedures but also to interpret contextual information, analyze problems, and make logical decisions. This reflects the principles of TKA, which emphasize reasoning and numeracy literacy as key components of academic ability (Kemendikdasmen, 2025). In this context, numeracy literacy extends beyond computational skills to include the ability to understand and apply mathematical concepts in real-life situations (Fajriah & Amalia, 2024). Therefore, the developed instrument aligns with the concept of authentic assessment, as it connects abstract mathematical knowledge with meaningful real-world contexts. This approach strengthens the role of assessment as both a measurement tool and a means of fostering students' cognitive development.

The analysis of students' response data further provides insight into the performance of the instrument. The average scores for each item fall within a relatively high range, indicating that most students were able to understand and respond to the items effectively. However, this finding requires careful interpretation. A high average score may suggest that the items are within an appropriate or moderate level of difficulty, but it may also indicate that the items are relatively easy for the tested group (Widaman & Revelle, 2023). From a measurement perspective, an optimal instrument should include a balanced range of difficulty levels to effectively differentiate students' abilities. Therefore, while the current instrument demonstrates good accessibility and clarity, further refinement is needed to enhance its discriminating power.

This interpretation is further supported by the relatively low standard deviation, which indicates limited variation in students' responses. On one hand, this suggests that the instrument has clear language and well-structured contexts, allowing students to interpret the items consistently. On the other hand, it may indicate that the instrument has not yet fully captured the diversity of students' abilities. This condition highlights the importance of incorporating items with varying levels of complexity, including more challenging problems that require deeper reasoning (Sun et al., 2025). Previous studies emphasize that while structured item development improves instrument quality, it must be complemented by advanced analysis such as item difficulty and discrimination indices to achieve optimal measurement effectiveness (Lambert & Newman, 2023; Swan et al., 2023). Thus, future development should aim to enrich the cognitive depth and variation of the instrument.

The findings of this study have important implications for mathematics learning evaluation. The developed TKA-equivalent assessment instrument can serve as an alternative tool for assessing students' mathematical reasoning more comprehensively. Its use enables teachers to move beyond procedural assessment toward evaluating deeper cognitive processes. Furthermore, integrating TKA-based instruments into classroom assessment can help familiarize students with reasoning-oriented and numeracy-based tasks from an early stage. This supports the shift toward competency-based education, where assessment is not merely an endpoint but an integral part of the learning (Sa'dijah et al., 2025; Zahroh et al., 2025; Zana et al., 2025). In this regard, the findings extend previous research by demonstrating that the integration of standardized frameworks such as TKA into instrument development can enhance both the quality and relevance of assessment practices. This also aligns with the broader goal of 21st-century education, which emphasizes critical thinking, adaptability, and problem-solving skills (Adeoye et al., 2024).

Despite these positive findings, this study has several limitations. The number of items developed is relatively small, consisting of only three questions, which limits the coverage of the full range of TKA indicators. Additionally, the sample was restricted to students from a single school, which may affect the generalizability of the results. Therefore, future research is recommended to develop a larger pool of items and involve more diverse samples to obtain more comprehensive and generalizable findings.

Overall, the results indicate that the developed TKA-equivalent mathematics assessment instrument meets the criteria of a good measurement tool in terms of validity and reliability. The instrument is capable of providing accurate and consistent information about students' mathematical reasoning abilities. These findings reinforce previous studies which

highlight that reasoning-based assessment instruments can offer a more comprehensive evaluation of students' academic abilities (Sa'dijah et al., 2025; Zahroh et al., 2025). Thus, this study contributes to the development of mathematics assessment practices that are aligned with TKA standards and oriented toward higher-order thinking skills.

CONCLUSION

This research produced a mathematical assessment instrument equivalent to the Academic Ability Test (TKA) for junior high school students developed through the ADDIE model. The instruments compiled based on the TKA grid succeeded in meeting the criteria of good measurement quality. The results of expert validation showed that the instrument was feasible to use after several improvements were made to the editorial aspect, clarity of the context of the questions, and the suitability of indicators. The empirical validity test shows that all question items have a corrected item-total correlation value that is greater than the r -value of the table so that it is declared valid. In addition, the results of the reliability test showed a Cronbach's Alpha value of 0.953 which was in the very high category so that the instrument had excellent internal consistency. These findings show that the assessment instruments developed are able to measure students' mathematical reasoning skills consistently and accurately. Thus, the assessment of TKA equivalents developed in this study can be used as an alternative evaluation instrument to measure the academic ability of junior high school students that are in line with TKA standards.

REFERENCES

- Adeoye, M. A., Prastikawati, E. F., & Abimbowo, Y. O. (2024). Empowering Learning: Pedagogical Strategies for Advancing 21st Century Skills and Quality Education. *Journal of Nonformal Education*, 10(1), 10–21. <https://doi.org/10.15294/jone.v10i1.1451>
- Adriantoni, Sari, W. S., Sari, N. P., & Fitri, Y. (2025). *Pemahaman Konsep Assesmen, Penilaian, Pengukuran, dan Evaluasi*.
- Aqidatul Aisyah, N., Anis Abdullah, A., Najib Mubarrok, M., Adawiya, R., Alin Sholihah, D., & Alma Ata, U. (n.d.). Penerapan Model Discovery Learning Berbasis Etnomatematika Berbantuan Geogebra terhadap Kemampuan Berpikir Kreatif Matematis. In *MATHEMA JOURNAL E-ISSN* (Vol. 6, Number 1).
- Fajriah, N., & Amalia, R. (2024). Pengembangan Modul Pembelajaran Berbasis Etnomatematika Masjid Sultan Suriansyah Pada Materi Bangun Ruang Sisi Datar Untuk Memunculkan Berpikir Kreatif Peserta Didik. In *Jurnal Mahasiswa Pendidikan Matematika* (Vol. 4, Number 2). <http://jtam.ulm.ac.id/index.php/jurmadikta>
- Hidayat, H., Mahmudah, H., Hendra, G., & Zaini, N. (2026). *Standar Penilaian sebagai Instrumen Evaluasi Kebijakan Pengelolaan Pendidikan Dasar*. 4, 2026. <https://doi.org/10.61104/alz.v4i1.4085>
- Jannah, S., Kholis, N., Sondang Sumbawati, M., & Achmad, F. (2026). *Pengaruh Keterampilan Berpikir Kritis Dan Berpikir Kreatif Peserta Didik Jurusan Teknik Instalasi Tenaga Listrik Dalam Menghadapi Tantangan Abad 21*.
- Kemendikdasmen. (2025). *Salinan Perkabang tentang Kerangka Asesmen TKA SD MI dan SMP MTs*.
- Lambert, L. S., & Newman, D. A. (2023). Construct Development and Validation in Three Practical Steps: Recommendations for Reviewers, Editors, and Authors.

- Organizational Research Methods*, 26(4), 574–607.
<https://doi.org/10.1177/10944281221115374>
- Muharromah, F., Fani, F. Al, & Ridho, A. R. (2025). *Reliabilitas Alat Ukur: Jenis-Jenis, Cara Pengukuran, Dan Faktor-Faktor Yang Mempengaruhinya Dalam Pembelajaran*.
- Munaroh, N. L. (2024). Asesmen dalam Pendidikan: Memahami Konsep, Fungsi dan Penerapannya. *Dewantara: Jurnal Pendidikan Sosial Humaniora*, 3(3), 281–297.
<https://doi.org/10.30640/dewantara.v3i3.2915>
- Paat, M., Armstrong F. Sompotan, Aneke Pesik, Yohanes Bery Mokal, & Imriani Moroki. (2024). Penerapan Multimedia Pembelajaran Berbasis Masalah dalam Meningkatkan Hasil Belajar dan Keterampilan Berpikir Tingkat Tinggi. *JURNAL PENDIDIKAN MIPA*, 14(1), 155–162. <https://doi.org/10.37630/jpm.v14i1.1493>
- Phafiandita, A. N., Permadani, A., Pradani, A. S., & Wahyudi, M. I. (2022). Urgensi Evaluasi Pembelajaran di Kelas. *JIRA: Jurnal Inovasi Dan Riset Akademik*, 3(2), 111–121. <https://doi.org/10.47387/jira.v3i2.262>
- Rahman, D. Y., Fajriah, N., & Suryaningsih, Y. (2021). Efektivitas Model Pembelajaran Core Terhadap Kemampuan Pemahaman Konsep Siswa Pada Mata Pelajaran Matematika di SMK. *JURMADIKTA*, 1(2), 11–20.
<https://doi.org/10.20527/jurmadijta.v1i2.794>
- Ramadhani, D. D. S., & Sukenti, D. (2023). Dampak Penerapan Model Problem Based Learning dalam Meningkatkan Kemampuan Berpikir Tingkat Tinggi Siswa. *Journal of Education Action Research*, 7(3), 327–335. <https://doi.org/10.23887/jear.v7i3.66698>
- Sa'dijah, C., Anwar, L., Utami, A. D., Dwi Cahyowati, E. T., Sulaiman, T., Laily Rahmawati, A. N., & Fitri, A. (2025). Integration of Technological Pedagogical and Content Knowledge in the Mathematics Learning Assessment Course: Strategies to Encourage Critical Thinking and Prepare Prospective Teachers to Face 21st-Century Challenges. *Journal of Physics: Conference Series*, 3148(1).
<https://doi.org/10.1088/1742-6596/3148/1/012004>
- Saputra, A., & Akmal. (2025). Pengembangan Instrumen Evaluasi (Analisis Pengembangan Penilaian Satuan Pendidikan MTs/SMP). In *Jurnal Pendidikan dan Keagamaan* (Vol. 2, Number 4).
- 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.
- Sun, J., Zheng, C., Xie, E., Liu, Z., Chu, R., Qiu, J., Xu, J., Ding, M., Li, H., Geng, M., Wu, Y., Wang, W., Chen, J., Yin, Z., Ren, X., Fu, J., He, J., Wu, Y., Liu, Q., ... Li, Z. (2025). A Survey of Reasoning with Foundation Models: Concepts, Methodologies, and Outlook. *ACM Computing Surveys*. <https://doi.org/10.1145/3729218>
- Swan, K., Speyer, R., Scharitzer, M., Farneti, D., Brown, T., Woisard, V., & Cordier, R. (2023). Measuring what matters in healthcare: a practical guide to psychometric principles and instrument development. *Frontiers in Psychology*, 14.
<https://doi.org/10.3389/fpsyg.2023.1225850>
- Syafi, M., Samsudin, M., Abidin, Z., Basarrudin, M., Tinggi Ilmu Tarbiyah, S., & Author, C. (2025). *EVALUASI PENDIDIKAN SEBAGAI DASAR PENGEMBANGAN INSTRUMEN PENILAIAN BERBASIS KOMPETENSI* ARTICLE INFO ABSTRAK (Vol. 1, Number 4). Oktober. <https://journal.yapakama.com/index.php/JAMED>
- Ulfah, M. (2025). INSTRUMEN PENGUJIAN PRODUK PEMBELAJARAN (PENGUJIAN VALIDITAS, PRAKTIKALITAS, EFEKTIVITAS). In *Jurnal Penelitian dan Pendidikan Agama Islam* (Vol. 3, Number 1).

- Utami, W. B., Setyosari, P., Sa'dijah, C., & Praherdhiono, H. (2025). Analysis of factors that influence student engagement in mathematics learning. *Perspectives of Science and Education*, 74(2), 257–269. <https://doi.org/10.32744/pse.2025.2.17>
- Widaman, K. F., & Revelle, W. (2023). Thinking thrice about sum scores, and then some more about measurement and analysis. *Behavior Research Methods*, 55(2), 788–806. <https://doi.org/10.3758/s13428-022-01849-w>
- Widiana, I. W. , G. I. K. , T. I. M. , & A. P. A. (2023). *Validasi penyusunan instrumen penelitian pendidikan*. . PT. RajaGrafindo Persada-Rajawali Pers.
- Zahroh, I. E. P., Anwar, L., & Chandra, T. D. (2025). *Assessment of Creative Thinking in Cultural Context in Junior High School Students*. <https://doi.org/10.35194/jp.v14i2.5192>
- Zana, F. M., Sa'dijah, C., Susiswo, Anwar, L., & Zulnaidi, H. (2025). *Development of assessment instruments aligned with curriculum: The need for an accurate measuring tool for high school students' critical thinking*. 030022. <https://doi.org/10.1063/5.0309129>
- Zayrin, A. A., Nopus, H., Maizia, K. K., Marsela, S., Hidayatullah, R., & Harmonedi, H. (2025). Analisis Instrumen Penelitian Pendidikan (Uji Validitas Dan Relibilitas Instrumen Penelitian). *Jurnal QOSIM Jurnal Pendidikan Sosial & Humaniora*, 3(2), 780–789. <https://doi.org/10.61104/jq.v3i2.1070>