

Development of an Evaluation Instrument for Attitude and Skills in Islamic Religious Education for Senior High School Students (Phase E)

Irwansyah ^{a.1,*}, Santi Lisnawati ^{b.2}, Wido Supraha ^{c.3}

^{*abc} Universitas Ibn Khaldun Bogor, Indonesia

^{*1} irsal_ges@yahoo.co.id ; ² santilisnawati@uika-bogor.ac.id ; ³ wido.supraha@uika-bogor.ac.id.

^{*}Correspondent Author

Received: 02-03-2026

Revised: 20-04-2026

Accepted: 25-06-2026

KEYWORDS

Islamic Religious Education;
Attitude Assessment;
Skill Assessment;
Instrument Validation.;

ABSTRACT

Evaluation in Islamic Religious Education (PAI) at the senior high school level should assess not only cognitive achievement but also students' attitudes and skills as part of authentic assessment in the Merdeka Curriculum. However, teachers often face difficulties in conducting objective affective and psychomotor assessments due to the lack of structured and validated instruments. This study aims to develop and examine the feasibility of an evaluation instrument for assessing attitudes and skills in Islamic Religious Education for Phase E senior high school students. This study employed a Research and Development approach using the ADDIE model. Data were collected through curriculum analysis, expert validation, and limited field testing. The development process included designing assessment indicators, constructing instrument items and rubrics, revising based on expert feedback, and conducting limited trials. The results show that the developed attitude and skills evaluation instruments are feasible for use. Expert validation results indicated a very good level of validity, while limited field trials showed that the attitude instrument was categorized as very good and the skills instrument as good. Therefore, the developed instruments can be used as structured and authentic tools to evaluate affective and psychomotor learning outcomes in Islamic Religious Education at the senior high school level.

This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Introduction

Islamic Religious Education (Pendidikan Agama Islam/PAI) plays a strategic role in the Indonesian education system in fostering students' faith, piety, and moral character through the balanced development of cognitive, affective, and psychomotor learning domains (Azra, 2020). At the senior high school level, the success of PAI learning is not solely reflected in students' mastery of religious knowledge, but also in the formation of religious attitudes and the development of practical religious skills that demonstrate the internalization of Islamic values in everyday life. This holistic orientation is consistent with Islamic teachings that emphasize the unity between faith, knowledge, and action, as stated in the Qur'an (Q.S. At-Taubah [9]: 105).

﴿وَقُلْ اَعْمَلُوا فَسَيَرَى اللّٰهُ عَمَلَكُمْ وَرَسُولُهُ وَالْمُؤْمِنُونَ﴾

which emphasizes that human deeds are observable and accountable before God, the Messenger, and society.

However, assessment practices in PAI remain predominantly focused on the cognitive domain, while the evaluation of attitudes and skills is often conducted through general observation without clear, measurable, and standardized instruments (Fadillah & Kurniawan, 2022). As a result, assessment outcomes tend to be subjective and less comprehensive (Wijaya, 2023). This condition is not fully aligned with the direction of the Merdeka Curriculum, which emphasizes holistic and authentic assessment to measure students' real competencies, including affective and psychomotor aspects (Mulyasa, 2021). Previous studies have highlighted the importance of systematic evaluation instruments in PAI learning (Fauzan & Anwar, 2022). Nevertheless, empirical research focusing on the development and validation of attitude and skills evaluation instruments for senior high school students within Phase E remains limited (Syarifuddin, 2023). This study adopts a research and development approach to produce a valid and practical evaluation instrument. Research and development is appropriate for educational assessment studies because it focuses on designing, testing, and refining products through systematic procedures (Ridwan & Syamsuddin, 2020). The development process in this study follows structured instructional design principles commonly applied in educational product development (Anafi et al., 2021). Furthermore, the structure and presentation of this article were adapted to international academic writing standards to ensure clarity, coherence, and consistency (Coussens et al., 2022). In the Indonesian education system, the Merdeka Curriculum emphasizes holistic learning outcomes that integrate knowledge, attitudes, and skills through authentic assessment practices. In Islamic Religious Education, this approach requires assessment instruments that are aligned with the expected learning outcomes at each phase, particularly Phase E at the senior high school level (Kemendikbudristek, 2022). Furthermore, assessment in the Merdeka Curriculum is designed to measure students' real competencies through continuous and meaningful evaluation processes, encouraging teachers to use various assessment techniques beyond written tests (Kemendikbudristek, 2023). In addition, the development of attitude assessment instruments in Islamic Religious Education requires clear indicators and appropriate scaling techniques to ensure accurate measurement of affective learning outcomes. Likert-scale instruments are widely used in affective assessment because they allow students to express the intensity of their attitudes in a structured and measurable manner (Sari & Anugrah, 2022). Furthermore, the assessment of religious skills in Islamic education must be grounded in authentic contexts that reflect students' real religious practices, enabling teachers to evaluate not only procedural accuracy but also consistency and meaning in students' actions (Rahman & Kholis, 2021).

Authentic Assessment and Evaluation of Attitudes and Skills in Islamic Religious Education

Authentic assessment emphasizes the measurement of students' real abilities through observable behaviors and meaningful learning activities. In Islamic Religious Education, this approach is essential to capture students' attitudes and skills as manifestations of internalized Islamic values, rather than merely assessing cognitive achievement (Zainuddin & Karim, 2021). In addition, the integration of digital media in Islamic Religious Education assessment supports the development of students' religious creativity, particularly through digital da'wah activities that reflect contemporary forms of religious expression (Zulkifli & Nasution, 2024). This principle aligns with the Qur'anic perspective that every individual will be judged based on their deeds (Q.S. Az-Zalzalah [99]: 7-8).

﴿فَمَنْ يَعْمَلْ مِثْقَالَ ذَرَّةٍ خَيْرًا يَرَهُ ﴿٧﴾ وَمَنْ يَعْمَلْ مِثْقَالَ ذَرَّةٍ شَرًّا يَرَهُ ﴿٨﴾﴾

which underlines the importance of evaluating real behavior and practice. Therefore, assessment in PAI should not only focus on what students know, but also on how they behave and apply Islamic teachings in daily life.

Evaluation of learning outcomes should be conducted systematically and comprehensively to ensure alignment between learning objectives, instructional processes, and assessment results. One evaluation framework that supports this holistic approach is the Context, Input, Process, and Product (CIPP) model, which evaluates educational programs by examining learning goals, instructional resources, learning implementation, and learning outcomes (Lisnawati, 2018). In the context of PAI learning, the CIPP model supports continuous improvement by enabling teachers to identify strengths and weaknesses in assessing students' attitudes and religious skills (Ulfah & Lisnawati, 2018).

Assessment of the affective domain in Islamic Religious Education includes students' attitudes, values, and moral behaviors that reflect Islamic teachings, such as honesty, discipline, responsibility, tolerance, cooperation, and self-reflection (Nurhayati & Abdullah, 2021). These values are central to Islamic education, as emphasized by the Prophet Muhammad ﷺ in his saying:

«إِنَّمَا بُعِثْتُ لِأَتَمِّمَ مَكَارِمَ الْأَخْلَاقِ»

(HR. Ahmad),

which highlights that the core mission of Islamic education is the formation of noble character.

Meanwhile, the psychomotor domain encompasses students' religious skills, including worship practices, Qur'anic literacy, social religious engagement, and the application of Islamic values in real-life contexts (Hidayat & Syafe'i, 2022). The Qur'an also emphasizes the importance of practicing religious teachings in real situations, as stated Q.S. Al-Māidah [5]: 48,

﴿لِكُلِّ جَعَلْنَا مِنْكُمْ شِرْعَةً وَمِنْهَاجًا﴾

which implies that religious guidance must be implemented in practical life.

Several studies indicate that the absence of clear behavioral indicators and standardized instruments often leads to subjective judgments in assessing affective and psychomotor domains in PAI learning (Fauzan & Anwar, 2022). To address this issue, evaluation instruments must be developed based on strong theoretical foundations and empirically tested to ensure their validity and reliability (Syarifuddin, 2023). Instruments grounded in educational and psychological theories, such as self-determination theory, can enhance students' intrinsic motivation and spiritual awareness by promoting reflective and meaningful assessment practices (Maulidya Nisa et al., 2024). Thus, the development of structured and validated evaluation instruments for attitudes and skills is not only pedagogically important but also normatively aligned with Islamic principles of accountability (muhasabah), practice ('amal), and character formation, making it essential for improving the quality of Islamic Religious Education learning at the senior high school level. Psychomotor assessment in religious education requires systematic observation supported by clear performance descriptors to reduce subjectivity and improve consistency of evaluation. Observation rubrics are considered effective tools for assessing psychomotor skills because they provide explicit criteria for each performance level, allowing teachers to conduct objective and structured assessments (Rahmawati & Wahyuni, 2023). In secondary education, psychomotor assessment should focus on students' ability to demonstrate skills through

repeated practice and real-life application, ensuring that learning outcomes are observable and measurable (Yusuf & Hamzah, 2020).

Method

This study employed a Research and Development (R&D) approach to develop and validate an evaluation instrument for assessing students' attitudes and skills in Islamic Religious Education (PAI) at the senior high school level (Phase E). The R&D approach was selected to ensure that the developed instrument met criteria of validity, practicality, and feasibility for classroom use (Sugiyono, 2020). The instrument development followed the ADDIE model, which includes five stages: Analysis, Design, Development, Implementation, and Evaluation (Anafi et al., 2021). At the analysis stage, curriculum review and teacher interviews were conducted to identify challenges in assessing affective and psychomotor learning outcomes within the Merdeka Curriculum for Phase E. The design stage focused on constructing instrument frameworks, indicators, and scoring rubrics. Two instruments were developed: a self-assessment questionnaire using a four-point Likert scale to measure students' attitudes, and a teacher observation rubric to assess students' skills. Instrument validity and reliability are essential to ensure that assessment results accurately reflect students' competencies. Validity testing ensures the relevance of indicators to learning objectives, while reliability testing ensures consistency of measurement results (Pratama & Fauzi, 2020).

During the development stage, the instruments were validated by experts in Islamic Religious Education, educational evaluation, and language to ensure content relevance, construct clarity, and readability. Revisions were made based on expert feedback prior to implementation. The implementation stage involved a limited field trial with Phase E senior high school students, while the evaluation stage examined the practicality and overall quality of the instruments based on trial results and teacher feedback. The final product is a valid and practical evaluation instrument for assessing attitudes and skills in Islamic Religious Education at the senior high school level. In addition, the inclusion of creativity indicators in the skills instrument is aligned with contemporary PAI learning approaches that emphasize project-based and contextual learning to foster students' creative expression of Islamic values (Setiawan & Mahmud, 2023).

Result

Needs Analysis for the Development of Attitude and Skills Evaluation Instruments

The needs analysis revealed that Islamic Religious Education (PAI) teachers at SMA Cahaya Rancamaya, SMA Pribadi Depok, and SMA Kharisma Bangsa experienced difficulties in conducting objective assessments of students' attitudes and skills. Assessment practices were still predominantly focused on cognitive aspects through written tests, while attitude and skills assessments were generally conducted through informal observations without standardized instruments, clear indicators, or detailed scoring rubrics. Consequently, assessment results were often subjective and difficult to document systematically.

Interview findings indicated that teachers required structured evaluation instruments that could support more objective, consistent, and well-documented assessment practices. Teachers also emphasized the need for practical instruments that could be easily implemented in daily classroom activities and aligned with the characteristics of Phase E senior high school students.

The questionnaire results further confirmed this need. The average level of need for developing attitude and skills evaluation instruments reached 91.2% at SMA Cahaya

Rancamaya, 92.0% at SMA Pribadi Depok, and 93.0% at SMA Kharisma Bangsa, indicating a very high demand across the three schools. Teachers strongly agreed on the importance of objective attitude assessment instruments, clear skills assessment rubrics, improved documentation of assessment results, practical assessment tools, and instruments that correspond to the learning characteristics of Phase E students.

Overall, the findings demonstrate a strong need for the development of structured and standardized attitude and skills evaluation instruments in Islamic Religious Education. These instruments are expected to support more objective, systematic, and sustainable assessment practices while facilitating the implementation of assessment principles within the Merdeka Curriculum.

Expert Validation Results

This section presents the results of expert validation and limited field testing of the developed evaluation instruments for attitudes and skills in Islamic Religious Education (PAI) at the senior high school level (Phase E).

Expert validation was conducted to examine the feasibility of the developed instruments in terms of content relevance, construct clarity, language appropriateness, and practical usability. Three experts were involved, consisting of an Islamic Religious Education content expert, an educational evaluation expert, and a language expert. The validation results are summarized in Table 1.

Table 1.1: Expert Validation Results of the Developed Instruments

Validator	Mean Score	Category
Content Expert (PAI)	3.70	Very Good
Evaluation Expert	3.60	Very Good
Language Expert	3.71	Very Good
Overall Mean	3.67	Very Good

The overall mean score of 3.67 indicates that the evaluation instruments are categorized as very good and are feasible for use without major revisions. Minor revisions suggested by the experts were related to wording clarity and indicator consistency.

Instrument Validity and Reliability Testing Results

Validity Testing

After expert validation and revision, the developed attitude and skills evaluation instruments were administered to 30 respondents to examine item validity. The validity analysis was conducted using the Pearson Product-Moment correlation technique with a significance level of 5% and a critical value of $r = 0.361$. An item was considered valid if its calculated correlation coefficient exceeded the critical value.

Table 1.2: Summary of Attitude Instrument Validity Results

Description	Result
Number of Items	32
Valid Items	30
Invalid Items	2
Invalid Item Codes	P03, P05
r-table ($N = 30, \alpha = 0.05$)	0.361
Range of Valid r-values	0.381 – 0.838

Source: Processed SPSS 31.0 data (2026).

As presented in Table 1.2, the attitude instrument initially consisted of 32 items. The validity analysis indicated that 30 items met the validity criterion ($r > 0.361$), while two items (P03 and P05) were excluded because their correlation coefficients were lower than the required threshold. The valid items produced correlation coefficients ranging from 0.381 to 0.838, indicating satisfactory item discrimination and a strong relationship with the overall attitude construct.

These findings demonstrate that the attitude instrument possesses adequate construct validity and is capable of measuring students' religious attitudes consistently. The removal of the two invalid items improved the overall quality of the instrument by ensuring that only representative indicators were retained. Consequently, the final attitude instrument consisted of 30 valid items that were considered suitable for assessing the attitudes of Phase E senior high school students in Islamic Religious Education.

The validity test of the skills instrument produced similarly positive results. All 20 items achieved correlation coefficients above the critical value of 0.361, indicating that every item satisfied the validity requirements.

Table 1.3: Summary of Skills Instrument Validity Results

Description	Result
Number of Items	20
Valid Items	20
Invalid Items	2
r-table ($N = 30, \alpha = 0.05$)	0.361
Range of Valid r-values	0.502 – 0.885

Source: Processed SPSS 31.0 data (2026).

As shown in Table 1.3, all items in the skills instrument were declared valid, with correlation coefficients ranging from 0.502 to 0.885. Several items demonstrated particularly strong correlations, such as Item 6 ($r = 0.842$), Item 9 ($r = 0.826$), and Item 11 ($r = 0.885$), indicating a strong contribution to the measurement of students' skills. These findings confirm that the instrument adequately represents the intended competency indicators and is capable of capturing variations in students' practical abilities. Therefore, all 20 items were retained in the final version of the instrument.

Reliability Testing

Reliability testing was conducted to determine the internal consistency of the developed instruments. Cronbach's Alpha was used as the reliability coefficient, with a minimum acceptable value of 0.60.

Table 1.4. Reliability Test Results

Instrument	Cronbach's Alpha	Minimum Standard	Category
Attitude Instrument	0.959	0.600	Very Reliable
Skills Instrument	0.951	0.600	Very Reliable

Source: Processed SPSS 31.0 data (2026).

As shown in Table 3, the attitude and skills instruments obtained Cronbach's Alpha values of 0.959 and 0.951, respectively. Both values are substantially higher than the minimum acceptable threshold of 0.60, indicating a very high level of reliability. According to commonly accepted reliability standards, coefficients above 0.90 are categorized as excellent, suggesting that the developed instruments possess outstanding internal consistency.

The high reliability coefficients indicate that the items within each instrument consistently measure the same underlying construct and are capable of producing stable measurement results. Therefore, the developed attitude and skills evaluation instruments can be considered reliable tools for assessing students' affective and psychomotor competencies in Islamic Religious Education. The combination of strong validity and excellent reliability provides empirical evidence that the instruments are suitable for educational assessment and practical implementation in Phase E senior high school learning environments.

Limited Field Trial Results of the Attitude and Skills Evaluation Instruments

The revised attitude and skills evaluation instruments were implemented in a limited field trial to examine their practicality, clarity, and descriptive measurement outcomes. The trial involved 192 Phase E senior high school students from several different schools, aiming to obtain representative data and reflect real classroom conditions in Islamic Religious Education (PAI) learning. The attitude evaluation instrument consisted of 30 items distributed across six dimensions, administered using a four-point Likert scale through student self-assessment. Meanwhile, the skills evaluation instrument comprised 20 indicators across five dimensions, assessed through direct teacher observation using a structured rubric with a scale of 1–4.

Results of the Attitude Evaluation Instrument

The descriptive analysis results show that all six attitude dimensions achieved mean scores above 3.00, indicating good to very good categories. The overall mean score of the attitude instrument was 3.29, categorized as very good, demonstrating that students generally exhibited positive religious attitudes in PAI learning.

Table 1.5: Recapitulation of Attitude Evaluation Results

Attitude Dimension	Attitude Dimension	Attitude Dimension
Faith and Piety	3.32	Very Good
Honesty	3.33	Very Good
Discipline and Responsibility	3.32	Very Good
Politeness and Care	3.29	Very Good
Tolerance and Cooperation	3.25	Good
Self-Reflection and Awareness	3.25	Good
Overall Mean	3.29	Very Good

The honesty dimension obtained the highest mean score (3.33), indicating strong integrity and academic honesty among students. The dimensions of faith and piety and discipline and responsibility also showed similarly high scores (3.32), reflecting students' consistency in religious observance and responsibility.

Meanwhile, the dimensions of tolerance and cooperation as well as self-reflection and awareness obtained slightly lower mean scores (3.25), although they still fell within the good category. This suggests that while students demonstrate positive social attitudes, reflective and collaborative skills require further strengthening through learning strategies that emphasize discussion, collaboration, and reflective practice.

Results of the Skills Evaluation Instrument

The skills evaluation instrument was tested on 10 students as an initial limited trial to assess the feasibility of observation-based assessment. The overall mean score of the skills instrument was 2.85, categorized as good, indicating that the rubric was practically applicable and capable of measuring students' religious skills objectively.

Table 1.6: Recapitulation of Skills Evaluation Results

Skills Dimension	Mean Score	Category
Worship Skills	2.95	Good
Qur'an Reading and Understanding	2.80	Good
Religious Social Skills	3.05	Good
Islamic and Reflective Thinking	2.75	Good
Religious Creativity	2.63	Good
Overall Mean	2.85	Good

The religious social skills dimension achieved the highest mean score (3.05), indicating that students actively participated in religious and social activities at school. In contrast, the religious creativity dimension obtained the lowest mean score (2.63), suggesting that students' creative expression in religious contexts, such as digital da'wah or project-based activities, remains limited.

Overall, the trial results indicate that both the attitude and skills evaluation instruments are understandable, practical, and feasible for use in assessing affective and psychomotor learning outcomes in Islamic Religious Education. The stability of scores across dimensions and the dominance of responses in the upper scale categories suggest that the developed instruments possess adequate descriptive measurement quality and are suitable for further implementation.

Discussion

The findings of this study demonstrate that the developed evaluation instruments for attitudes and skills in Islamic Religious Education (PAI) are feasible and effective for assessing affective and psychomotor learning outcomes among Phase E senior high school students. The overall results indicate that students' religious attitudes are in the very good category, while their religious skills are in the good category. These findings suggest that the developed instruments are capable of capturing students' learning outcomes systematically and providing meaningful information regarding their attitudes and skills in PAI learning.

The development of these instruments was initiated based on the results of the needs analysis, which revealed that teachers experienced difficulties in conducting objective assessments of attitudes and skills due to the absence of structured and standardized assessment instruments. In practice, assessments were often carried out through general observations without clear indicators, resulting in subjective judgments and inconsistencies in documenting students' learning outcomes. These findings indicate the need for valid, reliable, and practical assessment instruments that can support teachers in implementing more systematic evaluation processes in Islamic Religious Education.

To address these challenges, the study employed the ADDIE development model consisting of analysis, design, development, implementation, and evaluation stages. Through this systematic process, the instruments were developed based on the learning outcomes of the Merdeka Curriculum for Phase E students. The analysis stage identified assessment needs and curriculum requirements, while the design and development stages focused on constructing indicators, item statements, and observation rubrics. The implementation and evaluation stages ensured that the instruments underwent empirical testing and revision before being finalized. This systematic procedure contributed significantly to the quality and feasibility of the final product.

The results of the validity testing demonstrated that the developed instruments adequately measured the intended constructs. For the attitude instrument, 30 out of 32 items met the validity criteria and were retained in the final version, while two items were excluded

because they did not satisfy the required validity standards. Meanwhile, all 20 items of the skills instrument were declared valid. These findings indicate that the developed items successfully represented the dimensions of attitudes and skills targeted in Islamic Religious Education learning. The reliability testing further strengthened the quality of the instruments, with Cronbach's Alpha values of 0.959 for the attitude instrument and 0.951 for the skills instrument, indicating excellent internal consistency and stability of measurement.

Furthermore, expert validation conducted by subject-matter experts, educational evaluation experts, and language experts confirmed the overall quality of the developed instruments. The suggestions provided during the validation process were incorporated into the revision stage, resulting in improvements to indicator clarity, item wording, and assessment procedures. Consequently, the final instruments were considered valid, reliable, practical, and appropriate for classroom implementation.

The field-testing results further support the feasibility of the developed instruments. The attitude instrument achieved an overall mean score of 3.29, which falls within the "Very Good" category. This result indicates that the instrument is capable of systematically measuring students' religious and social attitudes and providing meaningful information regarding affective learning outcomes in Islamic Religious Education.

Among the assessed attitude dimensions, honesty obtained the highest mean score of 3.33, followed by faith and piety as well as discipline and responsibility, each with a mean score of 3.32. These findings suggest that students have developed positive attitudes toward honesty, religious commitment, and personal responsibility. The results indicate that Islamic Religious Education learning and the religious culture established within the school environment have contributed positively to the development of students' moral and spiritual character. Since one of the primary objectives of PAI is the internalization of Islamic values into daily life, the high scores in these dimensions reflect the successful implementation of value-based learning practices.

Although all attitude dimensions were categorized as good or very good, the dimensions of tolerance and cooperation as well as self-reflection and self-awareness obtained relatively lower mean scores of 3.25. Nevertheless, these dimensions remained within the good category. The findings indicate that students still require greater opportunities to develop collaborative and reflective competencies. Therefore, learning activities emphasizing cooperation, discussion, social interaction, and self-evaluation may contribute to strengthening these aspects of student development.

The skills assessment also demonstrated positive results, with the overall instrument categorized as good. These findings indicate that the observation-based rubric effectively measures students' religious skills and provides an authentic representation of their psychomotor competencies. As psychomotor learning outcomes are best assessed through direct observation of performance, the developed rubric aligns with the principles of authentic assessment promoted in the Merdeka Curriculum.

Among the skills dimensions, religious social skills achieved the highest mean score. This finding indicates that students actively participate in religious and social activities and are capable of interacting appropriately within religious contexts. The result suggests that school religious programs and activities play an important role in developing students' social competencies and strengthening the practical application of Islamic values in everyday interactions.

The worship skills dimension was also categorized as good, indicating that students generally perform ablution and prayer according to the required procedures. However, indicators related to concentration and sincerity in worship obtained relatively lower scores, suggesting that spiritual aspects of worship require continuous guidance and reinforcement. Similarly, the Qur'anic reading and comprehension dimension showed that students were generally able to read Qur'anic verses correctly, although their ability to explain and interpret the meaning of the verses remained relatively lower. These findings indicate that greater emphasis should be placed on developing students' understanding and internalization of Qur'anic teachings.

The Islamic and reflective thinking dimension also achieved a good category, indicating that students were able to relate Islamic values to daily situations. However, their ability to analyze problems and formulate solutions based on Islamic principles still requires further development. In addition, religious creativity obtained the lowest mean score among the skills dimensions, although it remained within the good category. This finding suggests that students are not yet fully accustomed to expressing Islamic values through creative products such as posters, presentations, videos, or other forms of educational media. Therefore, more innovative and project-based learning activities may be needed to strengthen students' creativity within Islamic Religious Education learning.

Overall, the findings confirm that the developed attitude and skills evaluation instruments meet the criteria of validity, reliability, and feasibility. The instruments provide teachers with structured assessment tools that can be used to evaluate students' affective and psychomotor competencies more objectively and systematically. In addition, the instruments generate valuable information regarding students' strengths and areas for improvement, enabling teachers to design appropriate follow-up programs. Therefore, the developed instruments can serve as practical tools for supporting the implementation of Islamic Religious Education assessment in accordance with the learning outcomes of the Merdeka Curriculum for Phase E senior high school students.

Conclusion

This study developed and validated attitude and skills evaluation instruments for Islamic Religious Education (PAI) for Phase E senior high school students using a Research and Development (R&D) approach based on the ADDIE model. The needs analysis involving PAI teachers from SMA Cahaya Rancamaya, SMA Pribadi Depok, and SMA Kharisma Bangsa revealed a strong demand for structured and objective assessment instruments, with a need level exceeding 91%. The developed instruments consist of six attitude dimensions measured through a four-point Likert-scale questionnaire and five skills dimensions assessed using a four-point observation rubric. Validity testing showed that 30 of 32 attitude items were valid, while all 20 skills items met the validity criteria. Reliability testing demonstrated excellent internal consistency, with Cronbach's Alpha values of 0.959 for the attitude instrument and 0.951 for the skills instrument.

The feasibility of the instruments was confirmed through expert validation, validity and reliability testing, and field trials. Expert validation resulted in an average score of 3.67 (Very Good), while field testing involving 192 students for the attitude instrument and 10 students for the skills instrument produced mean scores of 3.29 (Very Good) and 2.85 (Good), respectively. These findings indicate that the developed instruments are valid, reliable, practical, and feasible for assessing affective and psychomotor learning outcomes in Islamic Religious Education. Overall, the instruments provide teachers with structured and objective assessment tools that support the implementation of authentic assessment and the learning outcomes of the Merdeka Curriculum for Phase E senior high school students.

References

- Anafi, I., Suryadi, A., & Supriatna, N. (2021). Development of educational products using the ADDIE model: A systematic review. *Journal of Educational Research and Evaluation*, 5(2), 101–110. <https://doi.org/10.23887/jere.v5i2.34121>
- Coussens, A. H., Lee, J., & Smith, R. (2022). Writing and structuring research articles for international journals. *Journal of Scholarly Publishing*, 53(4), 215–229. <https://doi.org/10.3138/jsp.53.4.03>
- Fadillah, M., & Kurniawan, D. (2022). Authentic assessment in Islamic Religious Education: Challenges and practices in senior high schools. *Jurnal Pendidikan Islam*, 11(1), 45–58. <https://doi.org/10.14421/jpi.2022.11104>
- Fauzan, A., & Anwar, S. (2022). Pengembangan instrumen penilaian sikap dalam pembelajaran Pendidikan Agama Islam. *Tarbawi: Jurnal Pendidikan Islam*, 7(2), 133–145. <https://doi.org/10.32678/tarbawi.v7i2.5123>
- Kemendikbudristek. (2022). *Capaian pembelajaran Pendidikan Agama Islam dan Budi Pekerti SMA/SMK Fase E*. Jakarta: Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.
- Kemendikbudristek. (2023). *Panduan asesmen pada Kurikulum Merdeka*. Jakarta: Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.
- Lisnawati, S. (2018). Model evaluasi CIPP dalam pembelajaran pendidikan Islam. *Jurnal Evaluasi Pendidikan*, 9(1), 15–27.
- Maulidya Nisa, F., Rahman, A., & Hidayat, T. (2024). Self-determination theory in Islamic education assessment: Implications for student motivation. *International Journal of Islamic Educational Psychology*, 5(1), 22–36. <https://doi.org/10.18196/ijiep.v5i1.17892>
- Mulyasa, E. (2021). *Implementasi Kurikulum Merdeka dan asesmen autentik*. Bandung: Remaja Rosdakarya.
- Nurhayati, S., & Abdullah, I. (2021). Affective domain assessment in Islamic education learning. *Journal of Islamic Education Studies*, 6(2), 89–102. <https://doi.org/10.21580/jies.2021.6.2.6945>
- Rahmawati, L., & Wahyuni, S. (2023). Developing observation rubrics for psychomotor assessment in religious education. *Journal of Curriculum and Instruction*, 17(1), 55–68. <https://doi.org/10.17509/jci.v17i1.50211>
- Ridwan, M., & Syamsuddin. (2020). Research and development methods in educational assessment. *Journal of Educational Methodology*, 4(3), 211–220.
- Sugiyono. (2020). *Metode penelitian dan pengembangan (Research and Development)*. Bandung: Alfabeta.
- Syarifuddin. (2023). Evaluasi sikap dan keterampilan dalam pembelajaran PAI berbasis Kurikulum Merdeka. *Jurnal Pendidikan Agama Islam*, 20(2), 177–190. <https://doi.org/10.14421/jpai.2023.20206>
- Ulfah, M., & Lisnawati, S. (2018). Evaluasi pembelajaran Pendidikan Agama Islam menggunakan model CIPP. *Jurnal Pendidikan Islam*, 7(1), 33–46.

- Wijaya, R. (2023). Problematika penilaian afektif dan psikomotor dalam pembelajaran PAI di SMA. *Jurnal Studi Pendidikan Islam*, 6(1), 71–84.
- Zainuddin, M., & Karim, A. (2021). Authentic assessment implementation in Islamic education classrooms. *Al-Ta'lim Journal*, 28(2), 145–157. <https://doi.org/10.15548/jt.v28i2.674>
- Hidayat, T., & Syafe'i, I. (2022). Penguatan asesmen karakter dalam Pendidikan Agama Islam. *Jurnal Pendidikan Karakter*, 12(2), 201–214.
- Pratama, A., & Fauzi, M. (2020). Validity and reliability testing of educational assessment instruments. *Journal of Educational Measurement*, 4(1), 1–13.
- Rahman, F., & Kholis, N. (2021). Assessment of religious skills in Islamic education. *Journal of Islamic Learning*, 9(2), 119–130.
- Sari, D. P., & Anugrah, R. (2022). Development of Likert-scale instruments in affective assessment. *Journal of Educational Research*, 15(3), 245–258.
- Setiawan, A., & Mahmud, M. (2023). Project-based learning to enhance religious creativity in PAI. *Jurnal Inovasi Pendidikan Islam*, 4(2), 88–101.
- Yusuf, M., & Hamzah, A. (2020). Psychomotor assessment in secondary education. *International Journal of Educational Assessment*, 2(1), 33–47.
- Zulkifli, Z., & Nasution, S. (2024). Digital da'wah as religious creativity in Islamic education. *Journal of Islamic Education and Technology*, 3(1), 1–14.