

## **Training and Mentoring on Developing Critical Thinking Assessment Instruments for Islamic Civilization History Teachers in Madrasah Aliyah**

**Budi Santoso<sup>1\*</sup>, Mukhlas Triono<sup>2</sup>, Febrian Andi Hidayat<sup>3</sup>**

<sup>1</sup>Program Studi Pendidikan Pancasila dan Kewarganegaraan, Universitas Pendidikan Muhammadiyah Sorong, Papua Barat Daya

<sup>2</sup>Pendidikan Matematika, Universitas Pendidikan Muhammadiyah Sorong, Papua Barat Daya

<sup>3</sup>Pendidikan Ilmu Pengetahuan Alam, Universitas Pendidikan Muhammadiyah Sorong, Papua Barat Daya

\*Corresponding author: [budisantoso@unimudasorong.ac.id](mailto:budisantoso@unimudasorong.ac.id)

**Abstract:** The limited competence of Islamic Civilization History teachers in developing critical thinking assessment instruments remains a significant challenge in implementing the Merdeka Curriculum. Most teachers continue to use assessment practices that emphasize lower-order thinking skills and have limited understanding of the Facione critical thinking framework. This community service program aimed to strengthen teachers' competence through training and mentoring using the Participatory Action Learning (PAL) approach, which was selected to promote active participation, collaborative problem-solving, and continuous practice in authentic assessment development. The program involved 15 Islamic Civilization History teachers from six Madrasah Aliyah in Sorong Regency, Southwest Papua. Program effectiveness was evaluated through pre-test and post-test assessments, observations, and product evaluations. The results showed a substantial improvement in teachers' competence, with the mean score increasing from 61.33 to 87.13, representing an improvement of 42.07%. All participants successfully developed classroom-ready assessment packages consisting of assessment blueprints, HOTS-based test items, scoring rubrics, and assessment formats aligned with the six critical thinking dimensions proposed by Facione. Product evaluation further indicated that the developed instruments were relevant, practical, and suitable for classroom implementation. These findings demonstrate that the PAL-based training and mentoring model effectively enhances teachers' assessment literacy while producing validated assessment instruments that support authentic assessment practices and strengthen the implementation of the Merdeka Curriculum in Madrasah Aliyah.

**Keywords:** Facione framework; Islamic Civilization History; Teacher Professional Development.

© 2026 Samakta: Jurnal Pengabdian Kepada Masyarakat

**Received:** July 8, 2026

**Accepted:** July 11, 2026

**Published:** August 31, 2026

**DOI:** 10.61142/samakta.v3i2.445

**How to cite:** Santoso, B., Triono, M., & Hidayat, F. A. (2026). Training and Mentoring on Developing Critical Thinking Assessment Instruments for Islamic Civilization History Teachers in Madrasah Aliyah. *Samakta: Jurnal Pengabdian Kepada Masyarakat*, 3(2), 38–46. <https://doi.org/10.61142/samakta.v3i2.445>

*This is open access article under the CC-BY-SA license*



## INTRODUCTION

The implementation of the Merdeka Curriculum has shifted educational assessment from measuring factual knowledge toward evaluating higher-order thinking skills, particularly critical thinking. In this curriculum, assessment is expected to capture students' abilities to interpret information, analyze evidence, evaluate arguments, formulate logical conclusions, and solve contextual problems rather than merely recalling facts (Budi Santoso *et al.*, 2023; Santoso *et al.*, 2024). Consequently, teachers are required to possess adequate assessment literacy to design authentic assessment instruments capable of measuring complex cognitive competencies systematically and validly (UNESCO, 2025). Developing such instruments has become an essential component of teacher professionalism because assessment quality strongly influences the effectiveness of learning and the achievement of curriculum objectives.

Within Islamic Civilization History (*Sejarah Peradaban Islam*), the need for critical thinking assessment is particularly significant. The subject provides abundant opportunities for students to examine historical evidence, compare multiple historical perspectives, evaluate the contributions of Islamic civilization, and relate historical values to contemporary social issues (Istiningsih *et al.*, 2024). These learning characteristics are closely aligned with the six dimensions of critical thinking proposed by Facione—interpretation, analysis, evaluation, inference, explanation, and self-regulation (Facione & Gittens, 2016). Therefore, assessment in Islamic Civilization History should extend beyond factual recall and encourage students to construct historical reasoning, justify interpretations, and make evidence-based judgments.

Despite these expectations, many teachers continue to encounter difficulties in developing assessment instruments that effectively measure critical thinking (Calma & Davies, 2025); (Review, 2025). Preliminary needs assessments conducted through classroom observations, interviews, and discussions with Islamic Civilization History teachers from six Madrasah Aliyah in Sorong Regency revealed that assessment practices remain largely dominated by lower-order thinking skills (LOTS), emphasizing memorization of historical facts, chronology, and prominent figures. Teachers also reported limited experience in designing authentic assessment instruments and inadequate opportunities to participate in professional development related to critical thinking assessment. These findings indicate that strengthening teachers' assessment competence remains an urgent need to support effective implementation of the Merdeka Curriculum.

Previous community service and teacher professional development programs have demonstrated positive impacts on curriculum implementation, Higher Order Thinking Skills (HOTS), and assessment development (Agustin, 2025); (Barid, 2025). However, most existing programs primarily focus on developing HOTS-oriented questions or introducing general assessment concepts without systematically integrating Facione's critical thinking framework into assessment design. Furthermore, many initiatives conclude after short-term workshops and rarely provide continuous mentoring, peer review, expert validation, and opportunities for teachers to produce validated classroom-ready assessment instruments. Consequently, evidence regarding sustainable professional development models that combine conceptual understanding with intensive mentoring and authentic product development for Islamic Civilization History teachers remains limited.

This gap represents the principal novelty of the present community service program. Unlike previous initiatives, the program integrates the six dimensions of Facione's critical thinking framework with a Participatory Action Learning (PAL) approach that combines needs assessment, conceptual training, collaborative workshops, intensive mentoring, peer review, expert validation, and limited classroom implementation. Rather than merely improving teachers' theoretical understanding, the program emphasizes the production of validated

assessment packages consisting of assessment blueprints, HOTS-based test items, scoring rubrics, and assessment formats that are directly applicable in classroom practice. This integrated approach is expected to strengthen teachers' assessment literacy while providing a sustainable professional development model that can be replicated in other Madrasah Aliyah.

Therefore, this community service program aimed to improve the competence of Islamic Civilization History teachers in developing Facione-based critical thinking assessment instruments through training and mentoring using the Participatory Action Learning approach. Specifically, the program sought to (1) enhance teachers' understanding of critical thinking assessment principles, (2) strengthen their competence in developing authentic assessment instruments aligned with the Merdeka Curriculum, and (3) produce validated, classroom-ready assessment instruments that support meaningful learning and the implementation of authentic assessment practices in Madrasah Aliyah.

## METHOD

This community service program was conducted from May to June 2026 in Sorong Regency, Southwest Papua, Indonesia. The program involved 15 Islamic Civilization History (Sejarah Peradaban Islam) teachers representing six Madrasah Aliyah, namely MAS Muhammadiyah Aimas, MAS Nurul Yaqin Sorong, MAS Nurul Yaqin, MAS Integral Hidayatullah, MAN Persiapan, and MAS Al-Ma'arif. Participants were selected purposively based on a preliminary needs assessment indicating that they experienced difficulties in developing critical thinking assessment instruments aligned with the Merdeka Curriculum.

The program employed the Participatory Action Learning (PAL) approach because it emphasizes active participant involvement, collaborative problem-solving, experiential learning, and continuous reflection throughout the professional development process. Unlike conventional training that mainly focuses on knowledge transfer, PAL encourages participants to identify authentic classroom problems, collaboratively develop practical solutions, implement assessment products, and continuously improve them through mentoring and feedback. This approach was considered appropriate for strengthening teachers' assessment literacy while producing classroom-ready assessment instruments (Moleong & Surjaman, 2021); (Davies, 2015); (Santoso et al., 2025).

### **The implementation of the program consisted of five sequential stages.**

Stage 1: Needs Assessment and Preparation. Preliminary observations, interviews, and focus group discussions were conducted with school principals and participating teachers to identify existing assessment practices and professional development needs. The results were used to prepare the training modules, participant worksheets, evaluation instruments, mentoring schedules, and learning resources.

Stage 2: Conceptual Training. Participants attended interactive training sessions covering (1) authentic assessment in the Merdeka Curriculum, (2) assessment literacy, (3) Higher Order Thinking Skills (HOTS), (4) Facione's six critical thinking dimensions (interpretation, analysis, evaluation, inference, explanation, and self-regulation), and (5) principles of developing valid and reliable assessment instruments. Learning activities included presentations, case studies, discussions, and analysis of sample assessment instruments.

Stage 3: Workshop on Instrument Development. During the workshop, participants collaboratively developed critical thinking assessment instruments through several activities: analyzing learning outcomes (CP) and learning objectives (TP), formulating assessment indicators, preparing assessment blueprints, constructing HOTS-based test items, developing scoring rubrics, and compiling complete assessment packages aligned with the Facione

framework. Each participant produced an initial draft of an assessment instrument relevant to Islamic Civilization History learning.

Stage 4: Mentoring, Peer Review, and Expert Validation. Draft assessment instruments were reviewed through collaborative mentoring sessions involving facilitators and peer participants. Each instrument was evaluated for content relevance, alignment between learning objectives and assessment indicators, cognitive level of the test items, clarity of scoring rubrics, language appropriateness, and classroom applicability. Participants subsequently revised their instruments based on the recommendations received until the products met acceptable academic and pedagogical standards.

Stage 5: Limited Implementation and Program Evaluation. The revised assessment instruments were implemented on a limited scale in participants' respective schools to obtain initial feedback regarding their practicality and applicability. At the end of the program, participants reflected on their learning experiences and completed the program evaluation activities.

Program effectiveness was evaluated using four complementary instruments. First, pre-test and post-test assessments were administered to measure participants' understanding of critical thinking assessment concepts before and after the training program. Second, observation sheets were used to assess participants' engagement and performance during workshops and mentoring activities. Third, product evaluation rubrics were employed to evaluate the quality of the assessment instruments developed by participants, including assessment blueprints, HOTS-based test items, scoring rubrics, and assessment formats. Fourth, a participant satisfaction questionnaire was administered to obtain feedback regarding the relevance, usefulness, and implementation of the program.

The collected data were analyzed using descriptive statistical techniques. Pre-test and post-test scores were summarized using mean scores and percentage improvements to describe changes in participants' competence. Observation results, product evaluation scores, and participant responses were analyzed descriptively by calculating percentages and interpreting them according to predetermined evaluation criteria. Qualitative feedback obtained during mentoring sessions was used to complement the quantitative findings and provide a comprehensive evaluation of the program implementation.

The success of the community service program was determined using four indicators: (1) improvement in teachers' understanding of critical thinking assessment concepts, (2) enhancement of teachers' competence in developing Facione-based assessment instruments, (3) successful production of validated classroom-ready assessment packages, and (4) positive participant responses toward the training and mentoring activities. Achievement of these indicators was considered evidence that the PAL-based professional development model effectively strengthened teachers' assessment literacy and supported the implementation of authentic assessment in the Merdeka Curriculum.

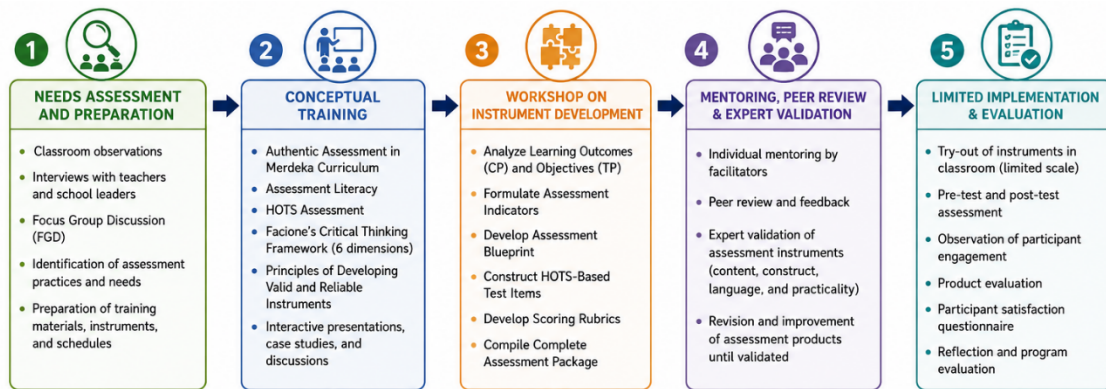


Figure 1. Implementation Process for Community Service Activities

## RESULTS AND DISCUSSION

### Implementation of the Training and Mentoring Program

The community service program was successfully implemented in collaboration with six Madrasah Aliyah in Sorong Regency, Southwest Papua, involving fifteen Islamic Civilization History teachers. The participating schools included MAS Muhammadiyah Aimas, MAS Nurul Yaqin Sorong, MAS Nurul Yaqin, MAS Integral Hidayatullah, MAN Persiapan, and MAS Al-Ma'arif. The program was conducted through five integrated stages consisting of needs assessment, conceptual training, workshops, mentoring and validation, and limited implementation.



Figure 2. Figure 1. Teachers Collaboratively Designing Critical Thinking Assessment Instruments

The initial needs assessment confirmed that although all participating schools had implemented the Merdeka Curriculum, most teachers still experienced difficulties in developing authentic assessment instruments capable of measuring students' critical thinking skills. Existing assessments predominantly measured factual knowledge, memorization of historical events, and comprehension of basic concepts, while tasks requiring interpretation, analysis, evaluation, inference, explanation, and self-regulation were rarely developed. These findings reinforced the urgency of strengthening teachers' assessment literacy through structured professional development.

During the conceptual training sessions, participants actively discussed authentic assessment principles, Higher Order Thinking Skills (HOTS), and the six critical thinking dimensions proposed by Facione. Interactive presentations, case studies, and collaborative discussions enabled teachers to critically examine their existing assessment practices and identify areas requiring improvement. The high level of participant engagement observed

throughout the training demonstrated that the learning activities were relevant to teachers' professional needs.

The workshop and mentoring sessions emphasized practical learning. Participants collaboratively analyzed learning outcomes and learning objectives before developing assessment indicators, assessment blueprints, HOTS-oriented test items, and scoring rubrics. Continuous feedback from facilitators and peer discussions encouraged iterative revisions, enabling participants to refine their assessment products until they met acceptable pedagogical standards. This collaborative process transformed teachers from passive recipients of information into active developers of authentic assessment instruments, consistent with the principles of Participatory Action Learning.

### Improvement of Teachers' Competence

Program effectiveness was evaluated using pre-test and post-test assessments administered before and after the implementation of the training and mentoring program. The evaluation was designed to determine the extent to which the program enhanced teachers' understanding of critical thinking assessment, authentic assessment principles, and the development of assessment instruments based on Facione's critical thinking framework. In addition to the written assessments, participants' progress was continuously monitored during the workshops and mentoring sessions through observations of their ability to formulate assessment indicators, construct Higher Order Thinking Skills (HOTS)-based assessment items, and develop appropriate scoring rubrics for Islamic Civilization History learning.

Overall, the evaluation indicated that the training and mentoring program had a positive impact on participants' professional competence. The results demonstrated improvements in teachers' conceptual understanding of critical thinking assessment as well as their practical ability to develop classroom-ready assessment instruments. Furthermore, the evaluation confirmed that all participants were able to complete comprehensive assessment packages aligned with the objectives of the Merdeka Curriculum and the six critical thinking dimensions proposed by Facione. A detailed comparison of participants' pre-test and post-test performance is presented in Figure 3.

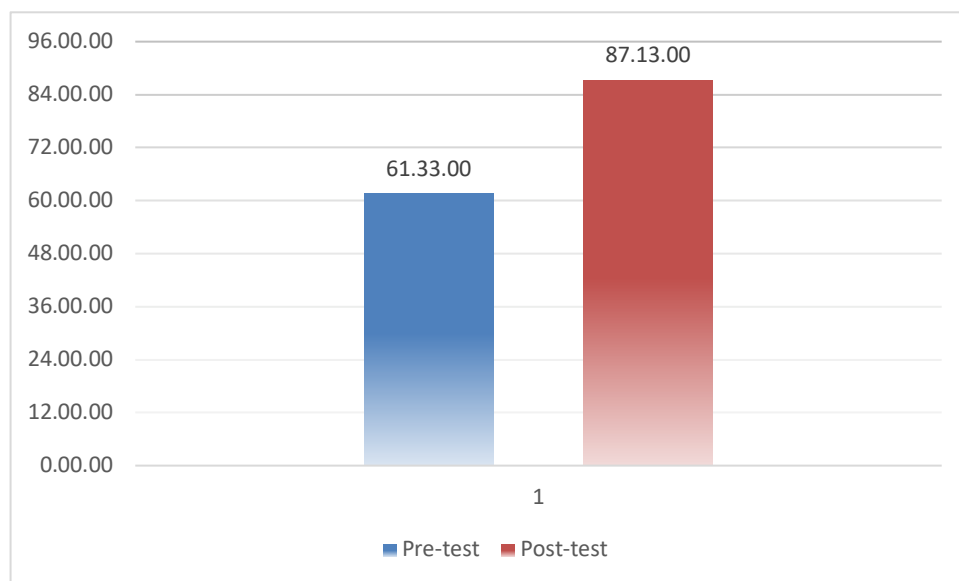


Figure 3. Average of participants' pre-test and post-test scores

The average score increased from 61.33 before the intervention to 87.13 after completion of the training, representing an improvement of 25.80 points (42.07%). More importantly,

every participant demonstrated higher post-test scores, indicating that the training benefited all teachers regardless of their initial competence level.

These findings indicate that combining conceptual understanding with intensive mentoring is more effective than one-time lecture-based training. Rather than merely introducing theoretical concepts, the PAL approach enabled teachers to apply newly acquired knowledge immediately through collaborative instrument development, continuous revision, and facilitator feedback. Such experiential learning strengthened both conceptual understanding and practical competence.

The observed improvement also supports adult learning theory, which emphasizes that professional competence develops more effectively when participants engage in authentic tasks, collaborate with peers, and receive continuous formative feedback. Consequently, the increase in post-test scores reflects not only improved knowledge but also enhanced confidence in independently developing authentic assessment instruments.

### **Development and Quality of Critical Thinking Assessment Instruments**

One of the principal outcomes of this program was the successful development of classroom-ready critical thinking assessment instruments. Every participant completed a comprehensive assessment package consisting of: assessment blueprint; learning indicators; HOTS-based essay questions; scoring rubrics; assessment format based on Facione's framework.

Unlike participants' previous assessment practices, which mainly measured factual recall and chronological understanding, the newly developed instruments required students to interpret historical evidence, analyze causal relationships, evaluate different historical perspectives, formulate reasoned conclusions, and relate Islamic civilization to contemporary issues.

Throughout the mentoring process, each instrument underwent iterative peer review and expert validation. Feedback focused on several aspects, including: alignment between learning outcomes and assessment indicators; cognitive level of assessment items; clarity of scoring rubrics; language appropriateness; classroom applicability.

Following revision, participants successfully improved the quality of their assessment products, producing instruments that were considered appropriate for classroom implementation. These findings indicate that continuous mentoring is essential for transforming theoretical understanding into practical assessment competence.

### **DISCUSSION**

The findings demonstrate that integrating conceptual training, collaborative workshops, and continuous mentoring effectively strengthened teachers' competence in developing critical thinking assessment instruments. Unlike conventional professional development activities that frequently emphasize theoretical knowledge, the PAL approach actively engaged teachers throughout the entire learning process. Participants were encouraged to identify classroom assessment problems, collaboratively develop solutions, revise their work based on expert feedback, and produce assessment instruments ready for implementation.

These findings are consistent with adult learning theory, which emphasizes that professional competence develops most effectively when learners actively participate in authentic tasks and continuously reflect upon their experiences. Collaborative mentoring further strengthened reflective practice by allowing participants to evaluate and improve their assessment products through peer discussions and facilitator guidance (Wahyuningsih *et al.*, 2026); (Santoso *et al.*, 2023).

The present findings also support previous studies reporting that sustained mentoring is more effective than one-off workshops in improving teacher competence. However, this program extends previous community service initiatives by integrating Facione's six critical thinking dimensions throughout the assessment development process. Consequently, teachers did not merely learn how to prepare HOTS questions but developed comprehensive assessment packages consisting of assessment blueprints, indicators, scoring rubrics, and classroom-ready instruments aligned with authentic assessment principles (Facione & Gittens, 2016).

From an educational assessment perspective, the resulting instruments represent valuable resources for implementing authentic assessment within Islamic Civilization History learning. The integration of interpretation, analysis, evaluation, inference, explanation, and self-regulation enables teachers to measure students' higher-order thinking skills more comprehensively than traditional factual assessments (Mahsun *et al.*, 2025). Furthermore, the availability of validated assessment instruments is expected to facilitate implementation of the Merdeka Curriculum while strengthening teachers' assessment literacy through sustainable professional learning communities such as MGMP.

Overall, these findings indicate that the PAL-based professional development model successfully achieved its objectives by improving teachers' competence while simultaneously producing validated assessment instruments ready for classroom implementation. The model therefore offers a practical and replicable strategy for strengthening assessment literacy among Islamic Civilization History teachers in other Madrasah Aliyah.

## CONCLUSION

This community service program successfully enhanced the competence of Islamic Civilization History teachers in developing critical thinking assessment instruments through a Participatory Action Learning (PAL) approach. The integration of conceptual training, collaborative workshops, intensive mentoring, peer review, and expert validation enabled participants to improve both their understanding of authentic assessment principles and their practical skills in designing assessment instruments aligned with Facione's six critical thinking dimensions. The substantial improvement in participants' pre-test and post-test scores further indicates that the program effectively strengthened teachers' assessment literacy.

An important outcome of the program was the successful production of classroom-ready assessment packages consisting of assessment blueprints, HOTS-based assessment items, scoring rubrics, and assessment formats that support authentic assessment practices within the Merdeka Curriculum. These products provide practical resources that can be directly implemented in Islamic Civilization History learning while encouraging teachers to integrate critical thinking assessment into their classroom practices more systematically.

The PAL-based professional development model demonstrated in this program offers a practical and replicable approach for improving teachers' assessment competence in Madrasah Aliyah. Future community service initiatives are recommended to involve a larger number of schools, incorporate classroom implementation on a broader scale, and evaluate the long-term impact of the developed assessment instruments on students' critical thinking skills and learning outcomes. Such efforts will contribute to strengthening authentic assessment practices and supporting the sustainable implementation of the Merdeka Curriculum.

## ACKNOWLEDGEMENTS

The authors would like to thank the principals and Islamic Civilization History teachers of the participating Madrasah Aliyah in Sorong Regency for their cooperation, enthusiasm, and active participation throughout the training and mentoring program. The authors also appreciate all colleagues and supporting staff who assisted in organizing and implementing the

activities. Their collaboration and commitment greatly contributed to the successful completion of this community service program and the development of critical thinking assessment instruments for Islamic Civilization History learning.

## REFERENCES

- Agustin, A. S. et al. (2025). *Development of Historical Thinking Skills Assessment Instruments*.
- Barid, S. S. A. Q. (2025). *A Systematic Review of Critical Thinking Assessment Instruments*.
- Budi Santoso, Triono, M., Ash Shiddiq, M. A., & Arifin, S. (2024). The Readiness of Islamic Religious Education Teachers to Enter The Era of Industrial Society 5.0. *Tafkir: Interdisciplinary Journal of Islamic Education*, 5(4), 624–636. <https://doi.org/10.31538/tijie.v5i4.1183>
- Calma, A., & Davies, M. (2025). Assessing students' critical thinking abilities via a systematic evaluation of essays. In *Studies in Higher Education*. <https://doi.org/10.1080/03075079.2025.2470969>
- Davies, M. (2015). *A Model of Critical Thinking in Higher Education* (pp. 41–92). [https://doi.org/10.1007/978-3-319-12835-1\\_2](https://doi.org/10.1007/978-3-319-12835-1_2)
- Facione, P., & Gittens, C. A. (2016). Think Critically. In *Insight assessment* (4th ed.). Pearson.
- Istiningsih, I., Unruh, T., Sutrisno, S., & Kurniawati, U. M. (2024). Active learning with research based on three principles of higher education. *Journal of Education and Learning*, 18(2), 421–429. <https://doi.org/10.11591/edulearn.v18i2.20966>
- Mahsun, Sumarmi, Utaya, S., Handoyo, B., & Wibowo, N. A. (2025). Enhancing Environmental Awareness: Evaluating the Impact of Project-Based Hybrid Learning on Critical Thinking for High School Students. *International Journal of Environmental Impacts*, 8(1), 123–135. <https://doi.org/10.18280/ijei.080113>
- Moleong, L. J., & Surjaman, T. (2021). Metodologi penelitian kualitatif. In *Remadja Karya* (p. 253).
- Review, P.-B. A. of C. T. A. S.-N. (2025). *Performance-Based Assessment of Critical Thinking: A Systematic-Narrative Review*.
- Santoso, B., In'am, A., Haris, A., & Wekke, I. S. (2024). Al-Islam and Kemuhammadiyah Learning Based on Religious Moderation in Multicultural Campus. *Al-Hayat: Journal of Islamic Education*, 8(1), 109. <https://doi.org/10.35723/ajie.v8i1.432>
- Santoso, B., Jaharudin, J., Rahman, A. A., & Triono, M. (2025). Community empowerment through the eco-mosque program at the Ali Bin Abi Tholib Mosque Sorong. *Journal of Community Service and Empowerment*, 6(1), 200–206. <https://doi.org/10.22219/jcse.v6i1.36601>
- Santoso, B., Triono, M., & Zulkifli, Z. (2023). Tantangan Pendidikan Islam Menuju Era Society 5.0: Urgensi Pengembangan Berpikir Kritis dalam Pembelajaran PAI di Sekolah Dasar. *Jurnal Papeda: Jurnal Publikasi Pendidikan Dasar*, 5(1), 54–61. <https://doi.org/10.36232/jurnalpendidikandasar.v5i1.2963>
- UNESCO. (2025). REIMAGINING OUR FUTURES TOGETHER: A new social contract for education. In *REIMAGINING OUR FUTURES TOGETHER: A New Social Contract for Education*. UNESCO. <https://doi.org/10.1163/9789004738058>
- Wahyuningsih, R., Muharom, F., & Annur, F. (2026). The Evaluation on Character Education in Indonesia in the Last Ten Years: Policy, Pedagogy, Management, and Assessment Perspectives. *Munaddhomah*, 7(2), 292–307. <https://doi.org/10.31538/munaddhomah.v7i2.2497>