

AI-RESILIENT PROCESS-BASED ASSESSMENT AND ORAL VERIFICATION IN S2 PAI UNZAH

Ahmad Zaini Al Ma'ruffi¹, Ainur Rofiq Sofa², Muhammad Ichsan³

^{1,2}Universitas Islam Zainul Hasan Genggong Probolinggo

³Universitas Hafshawaty Zainul Hasan

Email: ¹saikhzain2@gmail.com, ²bungaaklirik@gmail.com, ³ichsan29061997@gmail.com

Keywords:

Generative AI, Process-Based Assessment, Oral Verification, Academic Integrity, Graduate Islamic Education

Abstract

This article develops an AI-resilient process-based assessment and oral-verification model for the S2 PAI Program at Universitas Islam Zainul Hasan Genggong. Generative artificial intelligence can support idea generation, language revision, and feedback, but it also weakens the validity of unsupervised written products when student contribution cannot be verified. Using qualitative descriptive library research, the study synthesizes literature and policy guidance on generative AI, authentic assessment, academic integrity, oral assessment, and feedback. The findings show that assessment validity can be strengthened by collecting process evidence such as topic decisions, source notes, draft history, prompt disclosure, verification logs, feedback responses, and reflective justification. A short oral verification can probe core concepts, sources, methodological choices, and revisions. AI use should be classified by task purpose rather than treated through a universal ban, and students must disclose how tools affected the work. The study concludes that AI resilience depends on assessment redesign, clear rules, and verification of learning, not on unreliable detection alone. Equity, privacy, accessibility, and lecturer workload must be addressed in implementation.

INTRODUCTION

Generative artificial intelligence can rapidly produce essays, summaries, lesson plans, and research suggestions. These capabilities create opportunities for feedback and productivity, but they also challenge the validity of unsupervised written assessment. A polished product may no longer provide sufficient evidence that a student understands the concepts, evaluated sources, or made the decisions represented in the text (UNESCO, 2023).

Responses based only on prohibition and detection are fragile. AI detectors can produce uncertain results, while increasingly capable tools make authorship judgments difficult. Assessment should therefore focus on evidence of learning and responsible tool use. Process records, authentic application, and live explanation can make student thinking more visible without assuming that every use of AI is misconduct (TEQSA, 2023).

For S2 PAI, the issue includes academic integrity, Islamic ethics, source reliability, and the risk of fabricated religious references. Students need to verify Qur'anic, Hadith, classical, and scholarly sources rather than accepting generated claims. They also

need to distinguish language assistance from substitution of analysis and intellectual responsibility (Ramadani & Sofa, 2025).

This article aims to formulate an AI-resilient process-based assessment and oral-verification model for S2 PAI UNZAH. The study is conceptual and does not measure current AI use. It proposes assessment stages, disclosure categories, verification procedures, scoring principles, and safeguards for fairness and privacy.

LITERATURE REVIEW

Generative AI and assessment validity

Assessment validity concerns whether evidence supports the intended interpretation of student competence. When a task can be completed by an external system with minimal student judgment, the resulting product may be weak evidence of learning. The appropriate response is to redesign the relationship among outcomes, tasks, evidence, and verification (Cotton et al., 2023).

AI can still be educationally legitimate when the outcome includes critical use of digital tools. Students may compare generated explanations, identify errors, improve prompts, or document verification. The key is that permitted use must align with learning outcomes and be transparent to both students and assessors (UNESCO, 2023).

Process-based and authentic assessment

Process-based assessment collects evidence from stages of work rather than judging only the final product. Evidence may include problem selection, annotated sources, outlines, draft changes, peer feedback, methodological decisions, and reflection. This makes learning trajectories visible and supports formative intervention (Boud & Falchikov, 2006).

Authentic assessment further requires students to apply knowledge to contextual problems and audiences. A PAI student may analyze a local learning issue, design an intervention, or defend a research decision using verifiable sources. Context specificity reduces generic responses and strengthens the connection between assessment and professional practice (Nieminen et al., 2023).

Oral verification and academic integrity

Oral assessment allows lecturers to ask students to explain concepts, trace sources, justify choices, and respond to variations in the problem. A short verification does not need to repeat the entire written assessment. It samples high-value points where genuine understanding and authorship should be visible (Joughin, 1998).

Oral verification must be standardized enough to support fairness. Students should know the domains that may be questioned, assessors should use common prompts, and communication confidence should not be confused with conceptual mastery. Reasonable accommodations are necessary for disability, anxiety, language, and connectivity constraints.

RESEARCH METHODS

This study used a qualitative descriptive library-research method. Academic studies, international guidance, regulatory discussion papers, and literature on authentic assessment, oral assessment, feedback, and academic integrity were analyzed. Coding

focused on assessment validity, AI permissions, process evidence, oral verification, disclosure, equity, privacy, and implementation workload.

The concepts were synthesized into a staged assessment model for graduate PAI. No student submissions, AI-detection reports, or institutional cases were analyzed. The model therefore requires pilot testing, including evaluation of oral-verification reliability, student workload, accessibility, disclosure honesty, and the extent to which process evidence improves confidence in learning judgments.

RESULTS AND DISCUSSION

Classifying AI use by learning purpose

The first finding is that AI rules should be task-specific. A task may prohibit AI when independent recall or interpretation is the outcome; allow limited language support; require documented use for critique; or permit broad use when students remain responsible for verification and decision making. Each category should state permitted actions and required disclosure.

A universal ban may be unrealistic, while unrestricted use may invalidate the task. The classification should be explained in the assignment brief and rubric. Students should disclose the tool, purpose, representative prompts, outputs used, changes made, and verification procedures rather than submitting every interaction.

Collecting process evidence

Process evidence can include an approved topic statement, search strategy, annotated bibliography, source-verification table, outline, selected draft snapshots, feedback-response log, and final reflection. The aim is not continuous surveillance but a small set of meaningful checkpoints linked to outcomes.

For PAI assignments, students should verify religious quotations and scholarly references against authoritative sources. A verification log can record the generated claim, checked source, correction, and final decision. This practice turns the risk of AI hallucination into an opportunity for source criticism and academic honesty.

Conducting focused oral verification

A five- to fifteen-minute oral verification can ask students to summarize the argument, explain one key source, justify a methodological or pedagogical choice, and respond to a changed condition. Questions should be generated from the submitted work so that generic memorization is insufficient.

Scoring should emphasize conceptual understanding, evidence ownership, reasoning, and consistency with the written product. Students may consult the submitted work where appropriate. The oral result should complement, not automatically override, process and product evidence unless serious inconsistencies require further investigation.

Equity, privacy, and workload safeguards

AI-resilient assessment can create new inequities if some students have paid tools, stronger connectivity, or greater AI literacy. Requiring extensive prompt logs may expose personal data or create disproportionate administrative work. Oral assessment can also disadvantage students with communication anxiety or disability.

Implementation should provide access alternatives, data-minimization rules, accommodations, assessor training, and sampling strategies. High-stakes tasks may require oral verification for all students, while lower-stakes tasks can use targeted or random sampling. The purpose is credible learning evidence, not punitive surveillance (TEQSA, 2023).

Synthesis of discussion

The proposed model combines task-specific AI permissions, limited but meaningful process checkpoints, source verification, transparent disclosure, authentic contextual products, and focused oral questioning. These elements create overlapping evidence of learning. The model shifts academic integrity from detecting prohibited text toward demonstrating responsible authorship, verifiable reasoning, and accountability for decisions.

Table 1. AI-resilient assessment evidence framework

Assessment Stage	Evidence Collected	Integrity Contribution
Task briefing	AI-use category and rules	Clarifies permitted practice
Development process	Sources, drafts, prompt disclosure	Shows decision trajectory
Final product	Contextual scholarly output	Demonstrates application
Oral verification	Explanation and challenge response	Confirms understanding

The novelty of this article lies in integrating AI-use classification, source-verification logs, process checkpoints, and oral verification into one graduate PAI assessment design. It explicitly addresses the distinctive risk of fabricated religious references while avoiding dependence on AI detection. The framework treats responsible AI use as an assessable scholarly practice when aligned with learning outcomes.

CONCLUSION

This study concludes that AI-resilient assessment in S2 PAI UNZAH should be based on redesign rather than detection. Process evidence, task-specific permissions, disclosure, authentic context, source verification, and focused oral questioning can provide a stronger basis for judging student learning and academic responsibility.

Implementation must remain proportionate and equitable. The model should minimize unnecessary data collection, provide accommodations, ensure tool access alternatives, and manage lecturer workload. Future research should pilot the framework, compare judgment confidence with conventional assignments, analyze student disclosure practices, and examine the validity and reliability of short oral verification.

BIBLIOGRAPHY

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31(4), 399-413. <https://doi.org/10.1080/02602930600679050>

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315-1325. <https://doi.org/10.1080/02602938.2018.1463354>

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching*

- International, 61(2), 228-239. <https://doi.org/10.1080/14703297.2023.2190148>
- Joughin, G. (1998). Dimensions of oral assessment. *Assessment & Evaluation in Higher Education*, 23(4), 367-378. <https://doi.org/10.1080/0260293980230404>
- Khlaif, Z. N., Mousa, A., Hattab, M. K., Itmazi, J., Hassan, A. A., Sanmugam, M., & Ayyoub, A. (2025). Redesigning assessments for AI-enhanced learning: A framework for educational change. *Education Sciences*, 15(2), 174. <https://doi.org/10.3390/educsci15020174>
- Nieminen, J. H., Bearman, M., & Ajjawi, R. (2023). Designing the digital in authentic assessment: Is it fit for purpose? *Assessment & Evaluation in Higher Education*, 48(4), 529-543. <https://doi.org/10.1080/02602938.2022.2089627>
- Ramadani, S., & Sofa, A. R. (2025). Kejujuran dalam perspektif pendidikan Islam: Nilai fundamental, strategi implementasi, dan dampaknya terhadap pembentukan karakter santri di pesantren. *Jurnal Manajemen dan Pendidikan Agama Islam*, 3(1), 193-210.
- Sholihah, N., Sofa, A. R., Bukhori, I., Aziz, A., & Islam, M. H. (2026). Implementasi Canva sebagai media evaluasi interaktif dalam pembelajaran Pendidikan Agama Islam di kelas II SD Unggulan Cahaya Cendekia Glagah Pakuniran Probolinggo. *Al-Madrasah: Jurnal Ilmiah Pendidikan Madrasah Ibtidaiyah*, 10(2), 1154-1169.
- Sofa, A. R., Harifah, N., Khofifah, N., Habibah, W., & Karimah, U. N. (2025). Pengembangan penilaian pembelajaran berbasis VEO di MAN 2 Probolinggo. *JISPENDIORA: Jurnal Ilmu Sosial Pendidikan dan Humaniora*, 4(2), 503-511.
- TEQSA. (2023). Assessment reform for the age of artificial intelligence. Tertiary Education Quality and Standards Agency.
- UNESCO. (2023). Guidance for generative AI in education and research. UNESCO.
- Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research, and Evaluation*, 2, Article 2. <https://doi.org/10.7275/ffb1-mm19>
- Zhan, Y., Boud, D., & Dawson, P. (2025). Designing for authentic assessment: A scoping review. *Higher Education*. <https://doi.org/10.1007/s10734-025-01588-9>