



Vol. 4, No. 3, September 2026, pp. 2389-2407

INTERNATIONAL JOURNAL OF
POST AXIAL

FUTURISTIC TEACHING AND LEARNING

<https://journal.amorfati.id/index.php/postaxial> | ISSN 3025-7549



Beyond Correction: Revision, Re-examination, and Epistemic Validation in an Indonesian Pesantren

Muhammad Amruddin Latif^{1a}, Luqman Al Hakim^{2b}, Rifki Hamiyah Hadi^{3c}

¹²³Ma'had Aly At-Tarmasi, Pacitan, Indonesia

^alatifmuhammad1001@gmail.com, ^belhadeattarmasi@gmail.com, ^crifkihadi332@gmail.com

Article History:

Received:

12-08-2026

Revised:

22-08-2026

Accepted:

29-09-2026

Keywords:

Epistemic Practice; Validation Of Knowledge; Turās; Correction; Pesantren; Pondok Tremas

*Correspondence Address:

rifkihadi332@gmail.com

Abstract:

Research on pesantren has examined textual learning, memorization, legal deliberation, and assessment, but has paid less attention to what happens after a knowledge claim is corrected. This article investigates how correction, reworking, and re-examination shape epistemic validation across educational practices at Pondok Tremas. Drawing on a qualitative single-case study, it analyzes interviews, observations, and documents from Madrasah Aliyah, Ma'had Aly, Institut Agama Islam Attarmasi, and cross-institutional scholarly activities. Findings show that examination standards vary with the object of knowledge: kitab reading, Qur'anic memorization, legal formulations, and academic writing are assessed through different criteria. The relationship between correction and subsequent examination also varies. In *tahfiz* and the *risalah ilmiah*, revised work is returned for examination through identifiable procedures. In *baḥs al-masā'il*, formulations are examined collectively and may remain *mauqūf*. In *munāqasyah* and some course assignments, the available evidence does not consistently establish whether corrected work was re-examined. The article distinguishes correction, reworking, and re-examination as separate forms of epistemic work and shows why identifying an error alone does not establish that a revised result has gained recognition. These findings explain variation within one pesantren case without claiming that the same patterns characterize pesantren generally.

This is an open-access article under the [CC-BY-SA](#) license.



Introduction (مقدمة)

Knowledge in educational institutions is not merely transmitted; it is also presented, examined, and recognized through social practices. A reading may be judged accurate or mistaken, an answer may be accepted or rejected, and an argument may gain approval or remain

contested. These processes involve standards, actors authorized to conduct examinations, and procedures that determine the status of the outcome. Kelly and Licona define epistemic practices as socially organized, interactionally accomplished ways in which members of a group propose, communicate, evaluate, and legitimate knowledge claims (Kelly & Licona, 2018, pp. 139–140). This perspective shifts attention from knowledge as a finished product to the actions and social relations through which a claim gains, has yet to gain, or fails to gain recognition.

In pesantren education, knowledge is examined through diverse objects and activities. Reading a *kitab* requires accuracy in reading and explaining its meaning. Memorization requires correct pronunciation and sequencing, as well as the ability to retain memorized material. The discussion of legal problems requires relevant sources, reasons for applying them, and defensible formulations of a decision. Academic writing requires adequate alignment among the problem, sources, method, data, analysis, and conclusion. This diversity indicates that knowledge examination in pesantren cannot be understood solely through formal tests or assessment. Analysis must consider the object being examined, the standards applied, the actors who conduct the examination, and what happens after the result is communicated.

These practices take place within a tradition that provides texts, teachers, forms of training, and an understanding of proper performance. Asad describes a discursive tradition as discourse that instructs practitioners in the correct form and purpose of an established practice and relates it to the history through which that practice was formed (Asad, 1986, pp. 14–15). In his later work, Asad shows that language, bodily action, and repetition contribute to the acquisition of capacities within a tradition (Asad, 1986, pp. 14–15). In his later work, Asad shows that language, bodily action, and repetition contribute to the acquisition of capacities within a tradition (Asad, 2015, pp. 166–167). This framework situates examination practices in pesantren in relation to teaching, repetition, and the authorization of tradition. Asad, however, does not specifically explain how a claim moves from examination to legitimation. The discursive-tradition framework is therefore used to explain the context of practice, while the examination and recognition of claims are analyzed through the framework of epistemic practices.

Pesantren scholarship has provided important accounts of the transmission and transformation of these educational practices. Nurtawab identifies a shift in the status of *kitab*-based learning as traditional methods encounter modernization and demands for educational equivalence (Nurtawab, 2019). At the institutional level, recognition of pesantren within the national education system creates opportunities to strengthen traditional education while also introducing possible changes to long-reproduced values and practices (Nurtawab & Wahyudi, 2022). Gusmian and Abdullah show that manuscripts and *kyai-santri* networks are central to the transmission of multiple fields of knowledge in pesantren (Gusmian & Abdullah, 2022). These studies explain the continuity of tradition, methodological change, scholarly networks, and institutional restructuring. Their primary concern, however, is not how learning outcomes are examined and followed through until they acquire a particular status.

Research on assessment in pesantren comes closer to this problem. Puad and Ashton found that classroom assessment in the pesantren they studied was understood more in summative than formative terms (Puad & Ashton, 2021). Teachers gave considerable attention to grades, test results, conduct, and student accountability. These findings indicate that assessment in *pesantren* concerns not only academic ability but also relational structures and institutional expectations. Even so, an assessment-oriented account does not explain in detail how students respond to examination results, how work identified as weak is redone, or whether the revised result enters a subsequent examination.

The problem of examination also appears in studies of *bahās al-masā'il*. Riza shows that fatwas within Nahdlatul Ulama are formulated collectively through adherence to the *madhhab*

and efforts to respond to contemporary problems (Riza, 2011). In the context of Ma'had Aly, *uṣūl al-fiqh* is positioned as an important instrument for developing legal deliberation capable of addressing social problems (Juandi & Yasid, 2016). Zulkhairi et al. document the use of *ilhāq* and the contextualization of *turās* texts in *baḥs al-masā'il* forums (Zulkhairi et al., 2024). These studies explain sources, methods of reasoning, deliberative processes, and decision products. They do not, however, specifically distinguish among challenging an answer, revising a formulation in response to that challenge, and re-examining the revised formulation.

Studies of memorization and textual pedagogy reveal a different set of practices. Hardaker and Sabki treat orality, memorization, teacher-student interaction, and the study of sacred texts as interrelated elements of Islamic pedagogy (Hardaker & Sabki, 2015). Their subsequent research links memorization to the formation of capacities with physical and spiritual dimensions fisik dan spiritual (Hardaker & Sabki, 2016). Gent and Muhammad show that the relationship among memorization, recitation, and understanding cannot be assessed through the simplistic assumption that memorization without comprehension is necessarily meaningless (Gent & Muhammad, 2019). This literature clarifies the place of repetition, recitation, and memorization, but it does not focus on the sequence through which an error is identified, a reading or memorized passage is redone, and the new performance is resubmitted for examination.

The feedback literature shows why follow-up after examination requires separate attention. Hill and West demonstrate that dialogue about feedback can help students assess their work against criteria and formulate corrective action (Hill & West, 2022). The availability of feedback, however, does not ensure that students will use it in the same way. Xu et al. show that engagement with feedback is reflected not only in whether revision occurs but also in the depth, focus, and quality of the changes made (Xu et al., 2026). These findings underscore the need to distinguish the provision of feedback from the act of improving the work. Yet feedback scholarship has developed primarily in the contexts of writing and formal education. It has not been used to compare post-examination follow-up across *kitab* reading, memorization, legal deliberation, and academic writing within a single pesantren environment.

Existing studies explain how pesantren transmit knowledge, assess learning, and deliberate on legal questions. They offer less detail on what happens to a particular reading, answer, or written work after a weakness has been identified. In particular, correction is often discussed without distinguishing whether the work was subsequently changed and whether that changed result was examined again. This distinction matters because identifying an error, responding to it, and granting recognition to a revised result are different actions. Comparing these actions across several practices within one pesantren can clarify how examination leads to an accepted outcome or why it does not.

This article examines how correction, reworking, and re-examination relate to epistemic validation at Pondok Tremas. It compares *munāqasyah* and *takhassus* examinations, *tahfiz* recitation submissions, *baḥs al-masā'il*, papers and presentations, and the *risalah ilmiah* as practices within a single case. It asks: How do examination standards, correction, reworking, and re-examination interact in determining the status of knowledge claims across these practices??

Epistemic validation is used here as an analytical term for tracing the process through which a reading, answer, argument, or written work is examined against particular standards and is subsequently granted or not yet granted recognition. Correction refers to identifying an error, weakness, or element that requires improvement. Reworking refers to altering a reading, answer, formulation, or text in response to correction. Re-examination occurs when the reworked result is resubmitted to an authorized actor or forum, which may accept it, return it for further revision, or withhold approval. These three forms of work are analytically distinct and are not assumed to constitute a complete linear sequence in every case.

This article argues that epistemic validation at Pondok Tremas does not operate through a single uniform measure. Standards follow the object and form of work under examination, while correction, reworking, and re-examination are connected with varying degrees of traceability. Empirically, the article demonstrates this variation in direct examination, collective deliberation, and staged written accountability. Analytically, it clarifies the space between evaluation and legitimation within the framework of epistemic practices. Its contribution is not a typology of correction techniques, but a distinction among correction, reworking, and re-examination as epistemic work that may be connected without necessarily appearing as a complete sequence in the production of recognized knowledge.



Method (منهج)

This article presents a focused analysis of data generated in a broader qualitative case study of epistemic practices at Pondok Tremas. A qualitative approach was used because the analysis moved iteratively among data, evidence, and concepts to develop an increasingly close understanding of the phenomenon under study (Aspers & Corte, 2019, pp. 153–155). A case-study design was selected because validation practices cannot be separated from the institutional context, pedagogical relationships, knowledge sources, and authority structures in which they occur. Case studies permit in-depth examination of a contemporary phenomenon in its real-life context when the boundaries between phenomenon and context are not always clearly distinguishable (Yin, 2018, p. 15).

Pondok Tremas is treated as a single case. Madrasah Aliyah, Ma'had Aly, IAIT, and scholarly activities outside formal structures are treated as internal sites in which variation can be traced, not as independent institutional cases. Locating several internal sites within one case follows a single-case design with embedded units of analysis, provided that the analysis of each unit is ultimately related back to the case as a whole (Yin, 2018, pp. 51–54). Within this boundary, the article does not compare the performance of educational units. It examines variation in knowledge examination across *munāqasyah* and *takhassus* examinations, *tahfiz* recitation submissions, *baḥs al-masā'il*, papers and presentations, and the preparation of the *risalah ilmiah*.

The data derive from semi-structured interviews, observations, and documents. Participants in the broader study were selected purposively on the basis of their involvement, experience, position, and knowledge of the practices under investigation. Purposive selection requires a clear relationship between participant characteristics and the research purpose so that the relevance of the data can be assessed transparently (Campbell et al., 2020, pp. 654, 658). For this article, evidence was drawn from actors involved in conducting, supervising, examining, correcting, or accepting the outcomes of the analyzed practices. Semi-structured interviews retained a common set of core questions while allowing follow-up questions to elicit examples, reasons, sequences of action, and clearer contextual detail. This interview form balances a consistent focus with deeper exploration of participants' experiences (DeJonckheere & Vaughn, 2019, pp. 2, 5–6).

Observations focused on activities and interactions in which readings, answers, opinions, or written work were presented and examined. Attention was given to the object assessed, the examiner, the sources used, the standards stated or applied, the form of correction, and the responses that followed. Fieldnotes also retained the activity context, sequence of interaction, positions of the actors, and forms of verbal and nonverbal communication. Observational data were used primarily to distinguish normative provisions, participants' accounts, and actions visible in practice.

The documentary sources selected for this article included *munāqasyah* guidelines, *takhassus* examination materials, samples of student papers and a *risalah ilmiah*, *bahs al-masā'il* records, and the organizational structure of the Ma'had Aly deliberation forum. These documents were examined with attention to their content, function, and original purpose (Bowen, 2009, pp. 27–28, 30, 33). Guidelines and examination materials were used to identify prescribed standards; samples of written work were used to examine the form and content of academic submissions; and forum records were used to trace documented problems, sources, and decisions. Observational evidence concerning supervision, correction notes, and follow-up to revision was considered alongside these documents. The anonymized interview register and an inventory of the documentary and observational sources cited in this article are available through the online supplementary source (see the Data Availability Statement).

Data were selected for this article according to their relevance to epistemic validation. A data segment was included when it provided contextual information about at least one of the following: the object examined, the standard applied, the examiner, the form of correction, the response to correction, reworking, re-examination, or the status assigned after examination. Selection was not restricted to complete sequences. Evidence showing that correction was provided but that subsequent action could not be traced was retained because it defines the limits of the findings. An absence of evidence for reworking or re-examination was not interpreted as evidence that these actions never occurred. In such cases, the analysis states only that follow-up cannot be established from the available data.

Analysis involved repeated reading, coding, grouping, and comparison. First, relevant segments were reread in their original context so that their meaning remained connected to the activity, the actor's position, and the circumstances in which the data were generated. This step was necessary because coding can fragment data into separate elements and detach them from their broader context unless the original sources are revisited (Linneberg & Korsgaard, 2019, pp. 259, 267). Second, the data were coded for the object of examination, standards, examiner, form of correction, participant response, reworking, re-examination, and outcome status. Reworking was distinguished from a statement that a correction had been accepted: the code was used only when the data indicated a change to the reading, answer, formulation, or written work. Re-examination was likewise distinguished from reworking and was coded when the revised result returned to the attention of the assessing actor or forum.

Third, the data were compared at three levels. Cross-source comparison read interview accounts alongside relevant observations and documents. Comparison across actors identified differences between examiners and those whose work was examined. Comparison across practices identified commonalities and variation in the relationship among correction, reworking, and re-examination. The resulting empirical patterns were organized as direct examination, collective deliberation, and staged written accountability. Kelly and Licona's epistemic-practice framework was applied after these empirical patterns had been established to explain the relationship between evaluation and legitimation. The theory was not used to assume that every practice must follow the same sequence.

The credibility of the findings was examined by comparing interview accounts, observations, and documents where more than one type of evidence was available for the same practice. These sources did not provide equivalent coverage across all practices. Interview accounts were particularly important for tracing *tahfiz* correction and resubmission, while observations and documents provided different forms of evidence for examinations, deliberation, and academic writing. Divergent evidence and gaps in the available record were retained to delimit the conclusions (Korstjens & Moser, 2018, p. 2). Written provisions were not treated as evidence of implementation without comparison with participants' accounts or observed activities.

The researchers' positionality was considered throughout the analysis. The researchers have historical and professional proximity to Pondok Tremas as alumni and teaching staff. This proximity supported their understanding of terminology, activities, and institutional relationships, but also created a risk that familiar practices would be taken for granted and left unquestioned. Reflexivity is necessary to recognize how experience, assumptions, contextual position, and researcher-participant relationships shape the production and interpretation of data (Olmos-Vega et al., 2023, pp. 242, 244–245). This risk was managed by recording context and initial impressions, comparing data from actors in different positions, rereading documents and fieldnotes, and distinguishing normative provisions, participant accounts, and traceable practices. The conclusions are confined to Pondok Tremas as a single case and are not intended as statistical generalizations to all *pesantren*.

Result (نتائج)

The analysis shows that epistemic validation at Pondok Tremas does not follow a single uniform procedure. Variation concerns not only the type of activity, but also the object examined, the standards applied, how weaknesses are identified, what participants do after receiving correction, and whether the revised result is examined. The findings are therefore presented in four parts: object-specific standards, direct examination, collective deliberation, and staged written accountability.

Standards Follow the Object of Knowledge

In the Madrasah Aliyah *munāqasyah*, *santri* are asked to read an unvocalized text, provide a Javanese *pegon* gloss, explain the *murād* (intended meaning), and answer questions on *fiqh*, *naḥwu*, and *ṣarf*. Fluent recitation alone is therefore insufficient if a *santri* cannot explain the meaning or substance of the passage. Conversely, an apparently correct *fiqh* answer may still be challenged if the reading, gloss, or grammatical analysis is mistaken. Examiners use follow-up questions to move from reading to meaning, grammatical structure, and substantive mastery (MA-Doc-018; OBS-C).

Pegon glossing is not treated merely as a technique for translating the text. One of the *masyayikh* stated, "The reading must use *pegon* because that is how one masters *naḥwu* and *ṣarf*" (INT-A1-Masyayikh-02). The statement links word order, the grammatical position of each expression, and the meaning assigned to it. A *kitab* containing written glosses, however, does not establish that a *santri* can read and explain the text independently. Competence becomes visible only when the task is performed again before an examiner.

The Ma'had Aly *takhassus* examination has a similar structure but treats *fahm al-kitāb*, *fahm al-lughah*, active participation in study sessions, and completeness of the annotated *kitab* as separate assessment components. This separation distinguishes engagement in the learning process from the ability to read, interpret, explain, and analyze grammatical structure under examination conditions (MAT-Doc-011). Attendance and a complete *kitab* support the learning process but do not substitute for the competence that must be demonstrated in the examination.

Tahfiz applies different standards even though its object is also an Arabic text. Participants repeatedly referred to *makhraj*, *tajwid*, vowel length, correct verse sequence, fluency, the ability to recite without looking at the *muṣḥaf*, and the ability to retain memorization after an interval. Participants did not identify conceptual explanation of a verse's content as a condition for accepting a recitation submission. Memorization is assessed through the accurate, fluent, and stable reproduction of the text (INT-G-Tahfidz-01-05).

In *baḥs al-masā'il*, examination concerns the clarity of the problem description, the relevance of the *ibārah* (authoritative textual passage), the relationship between source and case, the reasons for applying the source, and the accuracy of the conclusion. Quoting a *kitab* does not by itself make an answer acceptable. Participants must identify the relevant passage and explain why the rule in the text applies to the problem under discussion. A lengthy quotation may be judged inadequate when its relationship to the problem is not explained, whereas a shorter quotation may suffice when its relevance is demonstrated clearly (INT-I-BM-01; MA-Doc-007-009; MAT-Doc-020-021).

For papers, presentations, and the *risalah ilmiah*, standards concern the adequacy of sources, conceptual clarity, the use of methods and data, the structure of the argument, and the relationship between the problem and the conclusion. A written product is available for inspection, but polished prose does not establish that the author understands the sources or can defend the argument. Presentations, supervision meetings, and examinations reopen the written text through questions about concepts, evidence, methods, data, and relationships among its parts (MAT-Doc-010; MAT-Doc-016; IAIT-Doc-024; OBS-C).

These findings show that similar forms of material do not produce identical standards. In *munāqasyah*, an Arabic text is examined through the relationship among reading, meaning, language, and *fiqh* content. In *tahfiz*, the text is examined through the accuracy and stability of oral reproduction. In *baḥs al-masā'il*, the text functions as a source whose relevance to the case must be explained. In academic work, sources form part of a sequence linking the problem, method, data, analysis, and conclusion.

Direct Examination and Correction in Performance

Direct examination shortens the interval between a participant's performance and the examiner's response. In *munāqasyah*, examiners provide prompts or follow-up questions when a reading is uncertain, a meaning is inaccurate, a grammatical analysis is incomplete, or a *fiqh* answer lacks a clear basis. Santri have an opportunity to recall and correct their answers during the session. *Munāqasyah* nevertheless remains an examination: prompts do not develop into extended instruction, and the final result is summarized as a grade (OBS-C).

The data do not show that every error in *munāqasyah* was followed by reworking and a further examination of the same passage after the session. A grade records the competence demonstrated during the examination, including responses to additional questions. It does not indicate whether an incorrect passage was subsequently restudied and re-examined. *Munāqasyah* therefore provides strong evidence of direct examination and outcome determination, but evidence of post-revision examination is uneven.

In *tahfiz*, correction is more closely embedded in the ongoing performance. A teacher or peer listener may tap, repeat the final word, or indicate the location of an error when the sequence, *makhraj*, *tajwid*, or vowel length is inaccurate. One participant explained that minor errors prompted "a tap or repetition of the last word," whereas more serious errors or excessive pauses could result in an instruction to *baleni*, or repeat the passage (INT-G-Tahfidz-04).

An indirect cue does not supply the correct continuation of the verse. The *santri* is first given an opportunity to recover the memorized passage. Another participant explained that if several prompts have been given and the recitation remains incorrect, the submission is rejected; the *santri* must step back, rehearse the passage until fluent, and then return to submit it again (INT-G-Tahfidz-05). Here, correction returns the participant to the same work: repeating the passage and producing a new performance for listening.

The *tahfiz* sequence reveals a closer relationship among correction, reworking, and resubmission. The data nevertheless do not record every participant's trajectory from an initial

error to final acceptance. The defensible finding is that the mechanism requires repetition and resubmission, not that every participant successfully completes the correction process.

Collective Deliberation and Withholding a Decision

In *baḥs al-masā'il*, validation begins before an answer is adopted. The problem description is first examined through requests for clarification about the actors, circumstances, terminology, or unclear aspects of the case. Clarification is necessary because an *ibārah* relevant to one set of circumstances may not apply to another that merely appears similar. Once the object is considered sufficiently clear, participants read the *ibārah*, propose an answer, and explain the relationship between source and problem (INT-I-BM-01; MAT-Doc-020).

Correction in the forum does not always take the form of declaring a participant's answer wrong. Participants or formulators may ask for the *ibārah* to be reread, question the relevance of a source, point out that an analogy still depends on an assumption, or request another basis. When deliberation stalls, the formulation team may separate aspects of the problem, identify key terms, or direct the search for references. Weakness is thereby located in the relationship among problem, source, reasoning, and conclusion rather than solely in the person who proposed the answer (INT-I-BM-01-02).

Re-examination occurs through forum deliberation. An answer may be maintained, revised, reformulated, or referred to the *mushahih* (verifier) for approval. The forum does not always reach a final decision. When the sources, information, or argumentative connections are judged inadequate, the matter may be designated *mauqūf* (suspended) (INT-I-BM-01). This status does not indicate that no epistemic work has occurred. It marks the forum's decision not to recognize any of the available answers.

The *baḥs al-masā'il* sequence differs from individual revision of a written text. Reworking may take the form of a jointly revised formulation, a search for additional sources, or a reconstructed conclusion involving participants, the formulation team, and the *mushahih*. The unit examined is therefore not only one participant's initial answer but a collective formulation that moves through the forum's division of labor. Records of the proceedings preserve the problem description, decision, and *ibārah* used, but do not always capture every objection and revision made before the decision was reached (MA-Doc-007-009; MAT-Doc-021).

Written Accountability and the Difference between Papers and the *Risalah*

During paper presentations, lecturers, *muhadlir*, or other participants may question conceptual clarity, the adequacy of data, the relevance of sources, and the relationship between the discussion and the conclusion. Authors respond in different ways: by explaining again, defending the argument, locating a passage in the paper, undertaking to revise it, or acknowledging that they cannot yet answer. These interactions expose the distance between a polished written text and the author's command of its content (OBS-C).

Corrections to a paper may be followed by requests to add sources, clarify concepts, improve the analysis, or adjust the conclusion. The data, however, contain more questions and directions than paired versions from before and after revision. Revised versions are not always reread through a uniform procedure. The evidence therefore supports the presence of correction and revision requests, but it is insufficient to claim that every paper was actually revised and systematically re-examined (MAT-Doc-010; IAIT-Doc-024; OBS-C).

The *risalah ilmiah* presents a longer sequence. Examination begins during supervision, when *mahasantri* are asked to clarify the problem, revisit a *kitab* or reference, explain how the method operates, add data or assess their adequacy, and align the analysis with the conclusion. After the

risalah examination, examiners provide notes on sections that still require revision. These notes return issues raised during the examination to the writing process (MAT-Doc-016; OBS-C).

Mahasnantri then revise the manuscript with reference to the examiners' notes. Revisions may replace or add sources, explain the operation of the method, reorganize the analysis, delimit a claim, or adjust the conclusion. The revised manuscript is resubmitted so that examiners can assess whether their earlier comments have been addressed. If the revisions remain inadequate, the relevant section may be returned for further work; if they are accepted, the examiner approves the revised result (OBS-C).

This traceability does not establish that all *risalah* are of equal quality or that every revision is successful on first resubmission. The finding is narrower: compared with an ordinary course paper, the *risalah* provides a clearer evidentiary trail linking supervision, examination, correction notes, textual revision, rereading, and examiner approval.

Variation in the Traceability of Validation

Comparison across practices shows that correction is the most readily identifiable element, while reworking and re-examination do not always leave equally clear traces. Table 1 summarizes this variation.

Practice	Object and primary standards	Form of correction	Reworking	Re-examination or outcome determination	Limits of the evidence
<i>Munāqasyah</i> and <i>takhassus</i>	Reading, meaning, <i>murād</i> , <i>fiqh</i> , <i>nahwu</i> , and <i>ṣarf</i>	Prompts and follow-up questions	May occur during the session	The outcome is determined primarily through grading	Post-examination review of each error is not uniformly traceable
<i>Tahfiz</i>	Makhraj, <i>tajwid</i> , vowel length, sequence, fluency, and retention	Cues, word repetition, prompts, or rejection of the submission	<i>Nderes</i> , <i>murāja'ah</i> , and repetition of the incorrect passage	The memorized passage is resubmitted	The data do not follow every participant from initial error to final acceptance
<i>Bahs al-masā'il</i>	Problem clarity, relevance of the <i>ibārah</i> , reasons for application, and conclusion	Clarification, objections, rereading of sources, and requests for another basis	Source searches and revision of the formulation	The formulation is accepted, revised, approved, or designated <i>mauqūf</i>	Reworking is collective and cannot always be attributed to one participant
Papers and presentations	Concepts, sources, data, argumentation, and explanatory ability	Questions, comments, and revision requests	May be requested after a presentation or assessment	Not always conducted through systematic rereading	Paired pre- and post-revision manuscripts are not uniformly available

<i>Risalah ilmiah</i>	Relationship among the problem, sources, method, data, analysis, and conclusion	Supervision, examination questions, and revision notes	The manuscript is revised in response to the notes	The revised version is reread before approval	Procedural traceability does not mean that all revisions or all <i>risalah</i> are of equal quality
-----------------------	---	--	--	---	---

Table 1 shows that correction, reworking, and re-examination are distinguishable across the practices studied. In *munāqasyah* and *takhassus*, the available evidence most clearly documents examination and outcome determination during the test; subsequent examination of corrected work is less consistently traceable. In *tahfiz*, an unaccepted passage is rehearsed and submitted again. In *bahs al-masā'il*, participants examine and may revise a collective formulation, while an unresolved matter can remain *mauqūf*. Papers and presentations generate questions and revision requests, but the resulting changes are not always documented or reread systematically. The *risalah ilmiah* provides a more identifiable route from examiners' notes to revision and subsequent review. These differences establish the empirical basis for examining how an evaluated result comes to be accepted, remains pending, or cannot be traced beyond the initial correction.



Discussion (مناقشة)

The findings show that examination standards vary with the object of knowledge and that correction does not always lead to a traceable revision and re-examination. This variation matters for epistemic validation because an identified weakness, a changed result, and recognition of that result are distinct events. The discussion therefore examines how the practices at Tremas connect evaluation to legitimation, while attending to the limits of what can be traced in each practice.

Validation Standards Follow Objects and Practices

The variation in standards at Tremas supports Kelly and Licona's account of epistemic practices as contextual. Forms of evidence, reasoning, representation, and justification depend on the domain and purpose of the activity (Kelly & Licona, 2018, pp. 140, 144–145, 156–157). Accurate *makhraj* and verse sequence are relevant to assessing a memorization submission but cannot assess the argument of a *risalah*. Conversely, adequate data and coherent analysis matter in academic work but do not substitute for accurate pronunciation in *tahfiz*. These differences do not constitute inconsistency when each standard is related to the object examined.

The findings extend the epistemic-practice framework in two directions. First, variation in practice occurs not only across communities or disciplines but also within a single educational environment. Second, the same source can perform different epistemic functions across activities. A *kitab* may be an object of reading, the basis for a legal answer, or a reference in academic writing. The authority of a source therefore does not by itself determine the adequacy of an answer. Adequacy depends on how the source is used and related to the object under discussion.

Mustofa et al. show that textual authority in pesantren *fiqh* reasoning is exercised through *qawli* and *ilhāqī* approaches as bases for legitimating decisions (Mustofa et al., 2024). The Tremas findings concur on the importance of textual authority but show that the presence of an *ibārah* does not complete the examination. The relationship among the wording of the source, the circumstances of the case, the reasons for application, and the conclusion must still be explained. Textual authority and the validity of textual use are therefore related but distinct issues.

Marwadi et al. show that pesantren use of *hisāb* brings together classical astronomical literature, Shafi'i *fiqh* legitimation, scholarly *sanad*, and calculation-based reasoning (Marwadi et al., 2025). Their study demonstrates that standards of justification may combine several types of sources and procedures according to the object. The Tremas findings add that these relationships among sources become visible in concrete work: reading, explaining, applying, comparing, and using them to support a conclusion.

In academic activities, the ability to locate sources must also be distinguished from the ability to use them. Bråten et al. show that views of epistemic justification influence the selection, processing, and use of multiple documents in written tasks (Bråten et al., 2022). This finding aligns with the Tremas data: a reference list does not establish an adequate argument if the author cannot identify the relevant passage and explain how the source supports the conclusion. Validation therefore extends beyond identifying source authority to accountability for the source's function in a particular work.

As a comparative lens, Longino's concept of public standards explains that criticism becomes relevant when it appeals to criteria recognized by participants in a community of inquiry (Longino, 2018, pp. 130–131). The concept is not used here as a normative checklist for measuring Tremas against the organization of scientific communities. Its relevance lies in the role of standards in preventing correction from becoming an inexplicable individual judgment. At Tremas, some standards are stated in examination guidelines and writing requirements. Others become visible through the questions asked, the errors identified, and the revisions requested. Variation in participants' understanding, however, shows that the existence of an institutional standard does not ensure a shared interpretation of it.

This finding complements Puad and Ashton's study, which identified a strong summative orientation among teachers in an Indonesian pesantren (Puad & Ashton, 2021). Their study highlights the roles of grades, test results, conduct, and student accountability. The Tremas data show that the summative-formative distinction does not fully explain the trajectory of an examined product. An assessment that ends in a grade may still include correction during the session, while an apparently formative comment does not necessarily lead to reworking followed by re-examination.

Validating Memorization as a Performance That Must Be Reproduced

The *tahfiz* literature has described memorization as a practice more complex than storing information. Gent identifies classroom routines and rhythms, student perceptions, the sacrifices involved in becoming a *hafiz*, and relationships between the *hifz* class and other areas of students' lives as major themes. Bensaid and Machouche locate Qur'an memorization within Muslim educational structures, authority, transmission, and the formation of character in traditional learning (Bensaid & Machouche, 2017). Both studies explain why memorization should not be treated as a learning technique detached from the educational and religious environment in which it is practiced.

Hardaker and Sabki connect orality, memorization, teacher-student interaction, and sacred texts within a holistic Islamic pedagogy (Hardaker & Sabki, 2015). Their subsequent study shows that memorization in madrasah life supports a process of personal embodiment with physical and spiritual dimensions (Hardaker & Sabki, 2016). Sabki and Hardaker further demonstrate that repeated reading, recitation, and writing support accuracy, fluency, and automaticity through iterative practice (Sabki & Hardaker, 2019). This literature explains the pedagogical function of repetition. The Tremas findings specify its evaluative function: errors are indicated through cues, the recitation is redone, and the result must be submitted again.

Gent and Muhammad caution that memorization without semantic understanding cannot simply be dismissed as meaningless because it is connected to broader forms of literacy, experience, and educational function (Gent & Muhammad, 2019). The Tremas findings support this caution. *Tahfiz* validation centers on the accuracy and stability of oral reproduction; conceptual explanation of a verse's content is not a condition for accepting a recitation submission. The present study adds, however, that acceptance is not determined by a participant's claim to have memorized the passage. Memorization must be reproduced as an audible performance that can be corrected, repeated, and resubmitted.

This addition matters because the memorization literature more often explains the functions of repetition, embodiment, identity, or the relationship between memorization and understanding. This article directs attention to the corrective trajectory within performance: how standards are applied during recitation and how a participant must produce a new performance after correction. Repetition thus serves both to form a capacity and to respond to a result that has not been accepted.

From Evaluation to Legitimation

Kelly and Licona distinguish the evaluation of claims from the process through which claims acquire legitimacy in a relevant community (Kelly & Licona, 2018, pp. 146–148). The Tremas findings clarify the relationship between these functions. Correction identifies a weakness in a reading, performance, formulation, or text; reworking changes that object; and re-examination allows an examiner or forum to judge the changed result. The distinction does not assume that every practice follows a complete sequence. It enables the analysis to identify where evaluation is followed by further work, where a result receives recognition, and where the available evidence does not establish what happened after correction.

Longino's concept of uptake helps explain the epistemic significance of such follow-up. Longino argues that criticism becomes part of a constructive justificatory practice when a community not only provides space for criticism but also responds by changing beliefs, data, reasons, or arguments (Longino, 2018, pp. 129–130). Used in this limited sense, correction at Tremas indicates that evaluation has occurred. Uptake becomes more visible when participants alter a reading, source, answer, formulation, or written text in response to correction.

This distinction is consistent with research on knowledge revision. Trevors et al. show that responses to refutation texts are influenced by how participants view authority, personal opinion, and the use of multiple sources as grounds for justification (Trevors et al., 2017). Correction therefore does not operate in a vacuum; participants' responses are shaped by their epistemic orientations. The present article does not measure the epistemic beliefs of Tremas participants. It examines the observable dimension: whether correction produces a change in the work being performed.

The feedback literature likewise confirms that providing comments is not equivalent to their use. Hill and West show that dialogic feed-forward becomes effective when students assess their work against criteria and formulate corrective action (Hill & West, 2022). Rastgou found that different forms of feedback generated different breadths of understanding concerning accuracy, content, organization, and fluency in writing (Rastgou, 2022). Both studies treat student response as integral to the process rather than as an automatic consequence of teacher comments.

Fitriyah et al. identify a gap between teachers' feedback practices and students' expectations concerning the type and relevance of feedback (Fitriyah et al., 2024). Students in particular regarded indirect feedback as less helpful in addressing specific problems in their writing. Their findings show that an available correction may not be recognized or used as its provider intended. The Tremas data do not specifically measure perceptions of feedback, but

varied responses to questions, admonitions, and written comments show that an opportunity to receive correction does not always become traceable reworking.

Lee shows that students may understand feedback as a report of results, the provision of a correct answer, or process information for improving learning (Lee, 2023). These differences are related to how students understand the function of assessment. The Tremas findings provide a practice-level counterpart: grades and outcome determinations do not perform the same function as corrections that direct participants back to their work. Analysis must therefore distinguish feedback that assigns a status to the result from feedback that initiates further work.

Xu et al. show that engagement with feedback is evident not only in whether revision occurs but also in the depth, focus, and quality of the changes. This article does not assess the quality of every revision made by Tremas participants (Xu et al., 2026). Findings from the *risalah* nevertheless show why re-examination is necessary: adding material does not necessarily address the initial problem, whereas narrowing a claim or replacing a source may constitute a more relevant change. A change must be assessed in relation to the correction, not by the amount of material added.

Yang et al. show that reflective assessment can support collective epistemic agency through monitoring, regulation of inquiry, negotiation, and synthesis of ideas (Yang et al., 2024). This finding is relevant to *bahṣ al-masā'il* despite the different context and technology. In the Tremas forum, the person who initially proposes an answer is not always the person who responds to correction. The search for sources, revision of the formulation, and determination of the result may be distributed among participants, the formulation team, and the *mushahih*. Uptake can therefore be individual or collective.

Collective Validation and the Possibility of Withholding a Decision

Studies of *bahṣ al-masā'il* generally emphasize sources, methods of reasoning, authority, and decision products. Riza shows that Nahdlatul Ulama fatwas are formulated collectively within a productive tension between adherence to the *madhhab* and the need to respond to contemporary circumstances (Riza, 2011). Juandi and Yasid demonstrate the central role of *uṣūl al-fiqh* in developing legal reasoning at Ma'had Aly (Juandi & Yasid, 2016). These studies situate *bahṣ al-masā'il* as institutional and methodological work rather than an unregulated exchange of opinions.

Zulkhairi et al. demonstrate the use of *ilhāq* and the contextualization of *turās* in deliberations on contemporary problems (Zulkhairi et al., 2024). Sulthon et al. show that *istinbāt al-aḥkām* can combine *qawli* and *manhajī* approaches (Sulthon et al., 2024). The Tremas findings do not offer a new classification of legal methods. Their contribution is to show what occurs when the relevance or application of a source is challenged: the answer must be re-explained, supported by another basis, revised, or withheld pending approval.

The forum's division of labor also has epistemic significance. Sobirin and Khasanah show that environmental theology and fatwas on global warming were discussed collaboratively by pesantren scholars and environmental scientists in a *halaqah* forum (Sobirin & Khasanah, 2023). The Tremas data show a comparable arrangement when expert information is required to clarify the object of a problem. Expert testimony does not replace the legal formulation, but it delimits the facts to be placed within a *fiqh* category. Validation in this context depends on a division of labor between explaining the object and determining its legal status.

Ibtihajuddin et al. identify dialogical clarification, ethical orientation, and socially embedded reasoning as important elements of *ijtihād irshādī* (Ibtihajuddin et al., 2025). Temuan Tremas menambahkan perhatian pada status rumusan sesudah klarifikasi dan The Tremas findings add attention to the status of a formulation after clarification and objection. A

formulation may be accepted, revised, approved, or left *mauqūf*. Re-examination need not return an assignment to an individual; it may occur as a formulation moves through several actors and stages of the forum.

Status *mauqūf* penting karena memperlihatkan bahwa validasi tidak selalu harus berakhir dengan The status of *mauqūf* is important because it shows that validation need not end in acceptance or rejection. Withholding a decision may acknowledge that the relationship among problem, source, and conclusion is not yet sufficiently strong. The absence of a final result does not necessarily indicate a failed forum. It may reflect epistemic discipline in refusing to confer recognition when the available basis remains inadequate.

Mufid shows that *baḥs al-masā'il* decisions on environmental issues may be followed through literacy and waste-management program (Mufid, 2020). Mustofa et al. likewise show that the contextualization of *fiqh al-mu'āmalah* can inform pesantren economic practices (Mustofa et al., 2024). Both studies concern action taken after a forum decision. The present article examines a different but related stage: how a formulation acquires a status that enables its use. Decision implementation follows the validation of the forum's product, whereas this article focuses on the internal route through which an answer moves toward acceptance or suspension.

Three Configurations of Validation

Comparison across practices reveals three temporal and organizational configurations: direct examination, collective examination, and staged written accountability. These configurations do not form a hierarchy from simple to more advanced practices. Each has a different temporal relationship, division of roles, and form of evidence.

In direct examination, performance and response occur almost simultaneously. This configuration permits errors in reading or memorization to be identified immediately. Two oral activities may nevertheless have different follow-up routes. *Munāqasyah* primarily determines the competence demonstrated in a single session, whereas *tahfiz* requires an unaccepted passage to be repeated and resubmitted. Temporal proximity between correction and response therefore does not determine whether the revised result will enter a subsequent examination.

In collective examination, validation work is distributed among several actors. An initial answer may move from a participant to the formulation team and the *mushahih* before acquiring a status. What is revised is not always an individual's work, but the formulation being developed by the forum. This configuration shows that the unit of uptake may be a collective claim rather than a change in one participant's understanding or writing.

In staged written accountability, the interval between correction and response is longer, but the trail of change can be preserved. Examiners' notes connect weaknesses in the manuscript to revision and subsequent rereading. The existence of a written text, however, does not automatically produce this sequence. A paper that receives comments but need not return to the examiner follows a different route from a *risalah* that must be revised before approval.

Traceability must not be equated directly with epistemic quality. Written practices materially leave more artifacts than oral interactions. A *risalah* is therefore easier to trace through manuscripts and revision notes, whereas oral correction may occur without producing an equivalent record. In this article, "more traceable" refers to the availability of evidentiary connections among correction, change, and outcome status. It does not rank the quality of practices.

Correction as Work within a Tradition

Asad explains that a discursive tradition teaches the correct form and purpose of an established practice. Argument is required when actors teach a practice, respond to doubt, or

address misunderstanding (Asad, 1986, pp. 14–16). On this account, correction need not be understood as an intervention from outside the tradition. It may be one of the means by which a tradition maintains accurate reading, examines source relevance, improves argument, and reworks inherited knowledge.

Asad also shows that language, repetition, and bodily action contribute to the formation of capacities within a tradition (Asad, 2015, pp. 166–167). The findings from *tahfiz* and reading examinations show that repetition does not merely reproduce an earlier form. It is directed by standards, opened to correction, and used to produce a more accurate performance. Continuity within a tradition may therefore involve reworking rather than only the transmission of an unchanged form.

Yaqin et al. show that connections between text and context, dialogue, and openness to difference form part of legal reasoning at Ma'had Aly (Yaqin et al., 2022). Yaqin et al. show that connections between text and context, dialogue, and openness to difference form part of legal reasoning at Ma'had Aly (Ishaq et al., 2025). This article does not propose another model of pesantren legal reform. Its explanatory addition operates at the level of micro-practice: the continuity of tradition is enacted through examining, correcting, revising, resubmitting, and granting or withholding recognition from a result.

This reading avoids two reductions of tradition. Tradition does not operate only through acceptance and repetition, because it also contains questions, corrections, objections, revisions, and suspended decisions. Correction, however, is not unregulated. It employs sources, standards, language, and divisions of authority recognized within the practice. Epistemic validation shows how continuity and reworking occur within the same tradition.

Contributions and Limits of Interpretation

This article offers three analytical contributions. First, validation is better understood as a relationship among actions than as a single assessment event. Second, correction, reworking, and re-examination must be distinguished because the presence of one does not establish the presence of the others. Third, the traceability of validation is shaped by temporal organization, the division of roles, and the material form of the work examined.

This interpretation remains subject to limitations. The data come from a single case and do not support a generalization that all pesantren use the same configurations of validation. Nor do the data follow every participant from an initial error to final acceptance. Paired pre- and post-revision manuscripts were not consistently available for course papers, while oral practices left more limited documentary traces. The article therefore does not assess the effectiveness of each practice, measure improvements in participants' competence, or infer that an absence of evidence for re-examination means that follow-up never occurred.

The defensible claim is not that one form of validation is superior to another. The contribution lies in showing that correction acquires epistemic significance through what follows it. When a corrected result is reworked and the change is examined, the relationship among standards, participant response, and outcome status becomes more traceable. When the evidentiary trail ends with a comment, grade, or untraceable revision request, the analysis can state only that validation proceeded as far as the evidence supports.



Conclusion (خاتمة)

This study shows that epistemic validation at Pondok Tremas varies according to the object of knowledge and the practice in which it is examined. *Kitab* reading, Qur'anic memorization, legal deliberation, and academic writing are assessed through different standards. Across these

practices, correction, reworking, and re-examination are distinct actions: identifying a weakness does not establish that the work was changed, and a change does not establish that the revised result was examined and accepted.

The relationship among these actions also varies. In *tahfiz*, an unaccepted passage is rehearsed and submitted again. In *baḥs al-masā'il*, a formulation is examined collectively and may remain *mauqūf*. The *risalah ilmiah* provides a clearer route from examiners' notes to revision and subsequent review. In *munāqasyah* and some course assignments, follow-up after correction is less consistently traceable. These findings clarify how evaluation may lead to recognition while showing that the full sequence cannot be assumed in every practice.

As a single-case study, the article does not establish whether the same patterns occur in other pesantren. Nor do the available data trace every individual correction through to final acceptance. Its contribution is to distinguish the actions between an initial judgment and the recognition of a revised result, within the limits of the evidence available.

Acknowledgment (شكر وتقدير)

We thank the leaders and administrators of Pondok Tremas, Madrasah Aliyah, Ma'had Aly At-Tarmasi, and Institut Agama Islam Attarmasi, as well as the teachers, *santri*, *mahasantri*, university students, and informants who contributed their time, access, and accounts during the research. We also thank those who assisted with observations and document retrieval. The authors remain responsible for all interpretations and limitations in this article.

Data Availability Statement

The anonymized interview register and the article-specific registers of documentary and observational sources are available at [Supplementary Source]. Interview transcripts, identifying information, and restricted institutional documents are not publicly available in order to protect participant confidentiality and comply with institutional access restrictions.



Bibliography (مراجع)

- Asad, T. (1986). *The Idea of an Anthropology of Islam* (Occasional Papers Series) [Occasional Paper]. Center for Contemporary Arab Studies, Georgetown University.
- Asad, T. (2015). Thinking About Tradition, Religion, and Politics in Egypt Today. *Critical Inquiry*, 42(1), 166–214. <https://doi.org/10.1086/683002>
- Aspers, P., & Corte, U. (2019). What Is Qualitative in Qualitative Research. *Qualitative Sociology*, 42(2). <https://doi.org/10.1007/s11133-019-9413-7>
- Bensaid, B., & Machouche, S. B. T. (2017). Memorizing the Words of God: Special Reference to 'Abdul Rahman Ibn Khaldun (D. 1406 A.D.). *Religious Education*, 112(4), 339–350. <https://doi.org/10.1080/00344087.2016.1224001>
- Bowen, G. A. (2009). Document Analysis as a Qualitative Research Method. *Qualitative Research Journal*, 9(2). <https://doi.org/10.3316/QRJ0902027>
- Bråten, I., Brandmo, C., Ferguson, L. E., & Strømsø, H. I. (2022). Epistemic justification in multiple document literacy: A refutation text intervention. *Contemporary Educational Psychology*, 71, 102122. <https://doi.org/10.1016/j.cedpsych.2022.102122>

- Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive Sampling: Complex or Simple? Research Case Examples. *Journal of Research in Nursing*, 25(8). <https://doi.org/10.1177/1744987120927206>
- DeJonckheere, M., & Vaughn, L. M. (2019). Semistructured Interviewing in Primary Care Research: A Balance of Relationship and Rigour. *Family Medicine and Community Health*, 7(2). <https://doi.org/10.1136/fmch-2018-000057>
- Fitriyah, I., Ningrum, A. S. B., & Gozali, I. (2024). An Investigation of Written Corrective Feedback in EFL Writing Assessment: How Teachers' Feedback Practices Meet Students' Expectations. *International Journal of Language Testing*, 14(1). <https://doi.org/10.22034/ijlt.2023.411616.1275>
- Gent, B. (2011). The world of the British *hifz* class student: Observations, findings and implications for education and further research. *British Journal of Religious Education*, 33(1), 3-15. <https://doi.org/10.1080/01416200.2011.523516>
- Gent, B., & Muhammad, A. (2019). Memorising and Reciting a Text without Understanding Its Meaning: A Multi-Faceted Consideration of This Practice with Particular Reference to the Qur'an. *Religions*, 10(7), 425. <https://doi.org/10.3390/rel10070425>
- Gusmian, I., & Abdullah, M. (2022). Knowledge Transmission and Kyai-Santri Network in Pesantren in Java Island during the 20th Century: A Study on Popongan Manuscript. *Afkar*, 24(1), 159-190. <https://doi.org/10.22452/afkar.vol24no1.5>
- Hardaker, G., & Sabki, A. (2016). The nature of memorisation for embodiment. *Journal for Multicultural Education*, 10(1), 87-98. <https://doi.org/10.1108/JME-01-2016-0019>
- Hardaker, G., & Sabki, A. A. (2015). Islamic pedagogy and embodiment: An anthropological study of a British Madrasah. *International Journal of Qualitative Studies in Education*, 28(8), 873-886. <https://doi.org/10.1080/09518398.2014.917738>
- Hill, J., & West, H. (2022). Dialogic Feed-Forward in Assessment: Pivotal to Learning but not Unproblematic. *Teaching and Learning Inquiry*, 10. <https://doi.org/10.20343/teachlearninginqu.10.20>
- Ibtihajuddin, Muh., Nur, I., Anshor, A. M., Mahfudz, A., & Mahmudah, R. (2025). Reconstructing Islamic Legal Epistemology: The Performative Logic of Ijtihād Irshādī in the Lajnah Bahtsul Masā'il of Pesantren Lirboyo. *Justicia Islamica*, 22(2), 461-488. <https://doi.org/10.21154/justicia.v22i2.10927>
- Ishaq, Faiz, M. F., & Ouich, A. (2025). Transcending without Disconnecting: The Integrative Epistemology of Ma'had Aly Situbondo and Its Model of Endogenous Islamic Legal Reform. *JURIS: Jurnal Ilmiah Syariah*, 24(2), 205-217. <https://doi.org/10.31958/juris.v24i2.15257>
- Juandi, W., & Yasid, A. (2016). Discourse of Islamic Jurisprudence in Indonesian Ma'had Aly between Taqlidy and Manhajy. *JOURNAL OF INDONESIAN ISLAM*, 10(1), 139. <https://doi.org/10.15642/JIIS.2016.10.1.139-158>
- Kelly, G. J., & Licon, P. (2018). Epistemic Practices and Science Education. In M. R. Matthews (Ed.), *History, Philosophy and Science Teaching: New Perspectives* (pp. 139-165). Springer International Publishing. https://doi.org/10.1007/978-3-319-62616-1_5
- Korstjens, I., & Moser, A. (2018). Series: Practical Guidance to Qualitative Research. Part 4: Trustworthiness and Publishing. *European Journal of General Practice*, 24(1), 1-6. <https://doi.org/10.1080/13814788.2017.1375092>
- Lee, M.-H. (2023). Lower secondary school students' perceptions of science assessment feedback and their conceptions of science assessment: A transitional phase of developing an understanding of assessment. *Research in Science & Technological Education*, 41(3), 928-945. <https://doi.org/10.1080/02635143.2021.1973401>
- Linneberg, M. S., & Korsgaard, S. (2019). Coding Qualitative Data: A Synthesis Guiding the Novice. *Qualitative Research Journal*, 19(3). <https://doi.org/10.1108/QRJ-12-2018-0012>
- Longino, H. E. (2018). *The Fate of Knowledge*. Princeton University Press.

- Marwadi, Niswah, E. M., Zain, M. F., Meidina, A. R., & Kafi, M. A. (2025). Traditional Islamic Legal Epistemology in Pesantren's Practice of Hīsāb. *Al-Ahkam*, 35(2), 429–458. <https://doi.org/10.21580/ahkam.2025.35.2.27687>
- Mufid, Moh. (2020). Green Fatwas in Bahtsul Masāil: Nahdlatul Ulama's Response to the Discourse on the Environmental Crisis in Indonesia. *AL-IHKAM: Jurnal Hukum & Pranata Sosial*, 15(2), 173–200. <https://doi.org/10.19105/al-lhkam.v15i2.3956>
- Mustofa, I., Hefni, W., Maliki, I. A., & Chamdan, U. (2024). Fiqh of Civilization at Traditional Islamic Boarding School: Contextualizing of Islamic Jurisprudence through Classical Texts of Mu'amalat. *JURIS: Jurnal Ilmiah Syariah*, 23(2), 351–365. <https://doi.org/10.31958/juris.v23i2.12245>
- Nurtawab, E. (2019). The Decline of Traditional Learning Methods in Changing Indonesia: Trends of Bandongan-Kitāb Readings in Pesantrens. *Studia Islamika*, 26(3), 511–541. <https://doi.org/10.36712/sdi.v26i3.11026>
- Nurtawab, E., & Wahyudi, D. (2022). Restructuring Traditional Islamic Education in Indonesia: Challenges for Pesantren Institution. *Studia Islamika*, 29(1), 55–81. <https://doi.org/10.36712/sdi.v29i1.17414>
- Olmos-Vega, F. M., Stalmeijer, R. E., Varpio, L., & Kahlke, R. (2023). A Practical Guide to Reflexivity in Qualitative Research: AMEE Guide No. 149. *Medical Teacher*, 45(3). <https://doi.org/10.1080/0142159X.2022.2057287>
- Puad, L. M. A. Z., & Ashton, K. (2021). Teachers' views on classroom-based assessment: An exploratory study at an Islamic boarding school in Indonesia. *Asia Pacific Journal of Education*, 41(2), 253–265. <https://doi.org/10.1080/02188791.2020.1761775>
- Rastgou, A. (2022). How feedback conditions broaden or constrain knowledge and perceptions about improvement in L2 writing: A 12-week exploratory study. *Assessing Writing*, 53, 100633. <https://doi.org/10.1016/j.asw.2022.100633>
- Riza, A. K. (2011). Contemporary Fatāwā of Nahdlatul Ulama: Between Observing the Madhhab and Adapting the Context. *Journal of Indonesian Islam*, 5(1), 35–65. <https://doi.org/10.15642/JIIS.2011.5.1.35-65>
- Sabki, A., & Hardaker, G. (2019). Memorisation and Learning in Sufi British Madrasahs. *Journal for Multicultural Education*, 13(4), 302–319. <https://doi.org/10.1108/JME-06-2019-0048>
- Sobirin, M., & Khasanah, K. (2023). The pesantren scholars' Fatwa on global warming and climate change: An integrative analysis of Islamic law, theology, and environmental sciences on the practice of multidisciplinary Ijtihad. *Cogent Arts & Humanities*, 10(1), 2193023. <https://doi.org/10.1080/23311983.2023.2193023>
- Sulthon, M., Syafi'i, I., & Nizami, A. G. (2024). Contemporary Fiqh in Indonesia: The Dynamics of Istibat al-Ahkām at Ma'had Aly Salafiyah Shafi'iyah Sukorejo Situbondo. *AHKAM: Jurnal Ilmu Syariah*, 24(1), 119–134. <https://doi.org/10.15408/ajis.v24i1.32174>
- Trevors, G. J., Kendeou, P., Bråten, I., & Braasch, J. L. G. (2017). Adolescents' epistemic profiles in the service of knowledge revision. *Contemporary Educational Psychology*, 49, 107–120. <https://doi.org/10.1016/j.cedpsych.2017.01.005>
- Xu, G., Li, D., & Gao, X. (Andy). (2026). Epistemic cognition and feedback engagement: A case study of Chinese L2 writers. *Assessing Writing*, 69, 101071. <https://doi.org/10.1016/j.asw.2026.101071>
- Yang, Y., Feng, X., Zhu, G., & Xie, K. (2024). Effects and mechanisms of analytics-assisted reflective assessment in fostering undergraduates' collective epistemic agency in computer-supported collaborative inquiry. *Journal of Computer Assisted Learning*, 40(3), 1098–1122. <https://doi.org/10.1111/jcal.12915>
- Yaqin, A., Wardi, M., & Mulyadi, A. (2022). Actualization of Moderation in Reasoning at Ma'had Aly Salafiyah Syafi'iyah Sukorejo Situbondo and Its Influence on Istibāth of Islamic Law. *Al-lhkam: Jurnal Hukum Dan Pranata Sosial*, 17(2), 434–457. Scopu. <https://doi.org/10.19105/al-lhkam.v17i2.5962>

- Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th ed.). SAGE Publications.
- Zulkhairi, T., Hajar, I., Safriadi, Marzuki, & Saifullah. (2024). Baḥsul Masāil at a Traditional Islamic Educational Institution in Aceh: Teungku Dayah's Contribution to the Development of Islamic Law. *Samarah*, 8(1), 579–601. <https://doi.org/10.22373/sjhk.v8i1.17408>