

FROM SOROGAN TO MAKNA GANDUL: ARABIC READING PEDAGOGY THROUGH KITAB KUNING IN INDONESIAN PESANTREN

Ammar Zainuddin¹, Hikmah²

¹Universitas KH. Abdul Chalim, Mojokerto, Indonesia; ²STAI Nurul
Hidayah, Riau, Indonesia

Corresponding author: Ammar Zainuddin, email: ammar@uac.ac.id

Article history: Received: 10 July 2026 | Revised: 9 August 2026 | Available
online: 30 August 2026

How to cite this article: Ammar Zainuddin & Hikmah. "From Sorogan to
Makna Gandul: Arabic Reading Pedagogy through Kitab Kuning in
Indonesian Pesantren." *Journal of Islamic Heritage and Civilization* vol. 2, no. 4
(2026): 569-604.

Abstract: Arabic reading instruction in Indonesian *pesantren* is characterised by pedagogical practices that support engagement with *kitab kuning*, including *sorogan*, *bandongan/wetonan*, *nahwu-sharaf* study, memorisation of grammatical *matn*, and vernacular glossing through *makna gandul* and *Pegon*. Previous studies have largely examined these practices separately, leaving their respective and interconnected contributions to Arabic reading competence and textual transmission insufficiently synthesised. This study examines how these practices support the reading, parsing, translation, comprehension, and interpretation of classical Arabic texts in *pesantren* contexts. Using a critical integrative review, it synthesises heterogeneous empirical, historical, and conceptual literature while distinguishing evidence by research design, context, and claim directness. The strongest evidence supports a recurrent pairing of *sorogan* and *bandongan* as individual verification and collective exposition, an association between *nahwu-sharaf* mastery and *kitab-reading* competence, and locally important roles for grammatical memorisation and *Pegon*-

based glossing. Alternative grammar methods and digital supplementation show, however, that Arabic reading pedagogy in *pesantren* is adaptive rather than methodologically uniform. The study therefore proposes a bounded transmission-ecology model in which textual materials, grammatical resources, teacher-mediated guidance, vernacular mediation, and adaptive media interact in developing Arabic reading competence while sustaining and reconfiguring locally embedded traditions of *kitab-kuning* literacy. The model identifies continuity-producing functions without assuming a single national system or deliberate institutional preservation.

Keywords: Arabic reading pedagogy; *Kitab kuning*; *Nahwu-Sharaf*; Textual transmission; Indonesian *pesantren*.

Introduction

Kitab-kuning literacy occupies a distinctive place in Indonesian *pesantren* because Arabic is learned not only as a language of communication but also as a medium for accessing legal, theological, exegetical, ethical, and grammatical scholarship. In many traditional and *salaf*-oriented settings, students encounter a repertoire that may include *sorogan* (individual guided reading before a teacher), *bandongan* or *wetonan* (collective teacher-led reading and explanation), explicit *nahwu-sharaf* (Arabic syntax and morphology) study, memorisation of grammatical *matn* (concise teaching texts), parsing, translation, and interlinear vernacular glossing. These practices are often discussed as markers of *pesantren* tradition, yet from the perspective of Arabic language education they also constitute techniques for developing access to complex written texts. Their pedagogical importance therefore lies not simply in their continuity as inherited routines, but in the specific reading functions they perform: directing attention to word endings, syntactic

relations, lexical meaning, textual coherence, teacher-sanctioned interpretation, and the gradual movement from supported to more independent engagement with Arabic scholarly prose. The present study takes this reading-oriented function as its disciplinary point of departure while treating *pesantren* textual traditions as the historical and institutional context in which those functions acquire meaning.

Arabic reading competence is understood here as a multidimensional capacity rather than a single score or universally standardised outcome. In the reviewed *pesantren* literature, competent kitab reading may involve accurate oral reading of minimally vocalised text, appropriate assignment of vowel endings, grammatical parsing, lexical and contextual translation, comprehension of propositions and relations across a passage, and explanation or interpretation that satisfies local criteria of scholarly adequacy. Different institutions may privilege these dimensions differently, and the available studies do not employ one common assessment instrument. This variation is analytically important. It prevents the review from treating 'reading ability' as a homogeneous outcome and makes it possible to ask which pedagogical practices support which dimensions of *mahārah al-qirā'ah*. *Sorogan* may foreground verification and correction; *bandongan* may model parsing and interpretation at scale; grammatical study may supply explicit linguistic resources; and *makna gandul* may externalise lexical, syntactic, and interpretive cues. These are hypotheses about function that must be bounded by the evidence rather than assumed as a universal instructional sequence. A case at Anwarul Huda Malang, for example, used direct book reading as a midterm Arabic assessment, illustrating how textual reading can function as an institutional criterion of competence rather than only as classroom practice (Jaunanto & Mahliatussikah, 2020).

The central problem is therefore not whether *pesantren* pedagogy is simply traditional or modern. That binary obscures differences in educational objectives and risks comparing practices designed for close reading of classical texts with pedagogies designed primarily for oral interaction or contemporary communicative performance. At the same time, continuity cannot be converted into proof of deliberate preservation. A method may persist because actors explicitly value an inherited form, but it may also persist through teacher formation, assessment routines, institutional path dependence, available resources, parental expectations, the social authority attached to particular modes of instruction, or combinations of these mechanisms. A rigorous account of Arabic reading pedagogy in *pesantren* must therefore distinguish observable instructional functions from the reasons institutions continue to reproduce them. This study treats alternative explanations as empirical possibilities rather than beginning with a predetermined claim that continuity necessarily equals conscious heritage preservation.

A second problem concerns the object called *kitab kuning*. Foundational bibliographic work demonstrates that the category refers to a historically varied corpus of books in Arabic script used in *pesantren* and related milieus, not to a single material format (van Bruinessen, 1990). Some texts are unvocalised, others are partially or fully vocalised, and the corpus includes concise *matn*, extended *sharh*, *hashiyah*, locally printed teaching texts, and multiple editorial formats. The reading demands of these materials are consequently not identical. *Kitab kuning* and *kitab gundul* overlap but are not synonyms, because lack of vocalisation describes a textual feature rather than the whole bibliographic tradition. Likewise, *makna gundul* is a glossing and reading practice, whereas *Pegon* is an Arabic-derived script used to write local languages; one can support the other without being analytically identical. These

distinctions matter for Arabic reading research because text format, script, glossing convention, and grammatical density shape what counts as difficulty, support, and competent performance.

The distinction also matters for the journal's broader problem of textual transmission. Transmitting a text is more than keeping a named teaching method alive. It includes the material text or edition, the linguistic code in which explanation occurs, the teacher who authorises or corrects a reading, the form of annotation, the institutional setting, and the assessment through which competence is recognised. Research on Arabic-script literary circulation in South and Southeast Asia shows that vernacularisation is not merely an auxiliary stage but a constitutive mode through which Arabic textual traditions are locally reworked (Ricci, 2011). Historical work on Javanese *makna-gandul* publications likewise demonstrates that interlinear translation can become a material publishing practice with local intellectual and cultural functions (Jamaluddin, 2019). From this perspective, Arabic reading pedagogy and textual transmission intersect: pedagogical routines organise how learners encounter texts, while the texts, scripts, glosses, and authorised readings constrain and shape the pedagogy itself. Recent manuscript research in *Senarai* likewise documents Arabic, Malay, and Maranao texts—including materials on Arabic morphology—as components of transregional Muslim knowledge networks in Southeast Asia (van der Molen, Katkova, & Wieringa, 2026).

The evidence base for these relations is heterogeneous and uneven. Some studies isolate *sorogan* and *bandongan*; others examine grammatical mastery, memorisation, *Pegon* training, or alternative grammar methods, while foundational scholarship concerns the material corpus and circulation of texts. The studies also differ in design, location, institutional type, and the dimension of reading competence they can support as evidence. This

fragmentation creates a problem of inference: co-presence, persistence, and functional complementarity cannot by themselves establish historical integration, causal effectiveness, or institutional intention. The present article therefore asks three bounded questions: (1) What documented functions do *sorogan*, *bandongan/wetonan*, grammatical memorisation, and vernacular glossing perform in the development of Arabic reading competence for *kitab-kuning* texts? (2) To what extent does the literature support functional coordination among these practices, and where do alternative or partial pathways to reading competence appear? (3) What can the evidence support concerning textual continuity, teacher authority, adaptation, and constraint without inferring a uniform national system or deliberate preservation from persistence alone?

The contribution is an evidence-bounded transmission-ecology model of Arabic reading pedagogy. The term ecology is used analytically rather than historically: it describes recurrent relations among textual objects, grammatical resources, modes of teacher-mediated verification and exposition, vernacular inscription, and changing instructional media. Its novelty lies in connecting these elements without converting them into a fixed sequence or assuming that every *pesantren* reproduces the same arrangement. The model also keeps two levels of analysis in view. At the pedagogical level, it asks how combinations of linguistic knowledge, guided reading, correction, glossing, and adaptive media shape particular dimensions of Arabic reading competence. At the textual-transmission level, it asks how those pedagogical arrangements contribute to the continued circulation, interpretation, and local re-inscription of *kitab-kuning* texts. This narrower formulation makes the model testable against disconfirming cases and aligns the article's contribution to Arabic language education with a historically and textually

grounded inquiry into Indonesian Islamic intellectual transmission.

Literature Review

1. *Kitab Kuning, Kitab Gundul, and the Material Text*

Van Bruinessen's (1990) mapping of books in Arabic script used in the *pesantren* milieu provides a necessary starting point because it treats *kitab kuning* as a corpus with disciplinary, bibliographic, and historical variation. The implication for pedagogy is substantial: the reading demands of a concise *matn*, an extended *sharh*, a *hashiyah*, or a locally glossed edition are not materially identical. A learner reading a short grammatical primer may be expected to recover dense rules from compact wording, whereas a learner reading a legal or theological commentary may need to track argument, quotation, embedded commentary, and technical vocabulary across a much longer passage. Nor is lack of vocalisation a sufficient definition of the category. In this article, *kitab kuning* denotes the broader *pesantren*-associated corpus, while *kitab gundul* is reserved for texts or editions whose Arabic is substantially unvocalised. This distinction keeps the pedagogical object precise: Arabic reading competence must be interpreted in relation to the material and linguistic features of the text actually being read rather than inferred from the label *kitab kuning* alone. Research with Arabic-speaking readers also shows that performance differs across voweled and unvoweled forms and is affected by linguistic distance between spoken and standard varieties (Saiegh-Haddad & Schiff, 2016). Although Indonesian *pesantren* learners inhabit a different linguistic ecology, this evidence reinforces the need to treat vocalisation and text form as substantive reading variables rather than incidental features.

Material form also changes what counts as transmission. A printed *kitab* can carry authorial text,

editorial arrangement, marginal commentary, vocalisation, and handwritten student annotations simultaneously. The pedagogical event may add oral translation, syntactic parsing, teacher commentary, locally conventional signs, and interlinear glosses. These layers make the reading act both linguistic and textual: students do not encounter abstract Arabic sentences detached from a tradition of use but texts embedded in particular layouts, commentarial relations, and authorised reading routines. Dhofier's (2011) account of *pesantren* scholarship helps situate teacher authority and textual mastery within a wider institutional culture, while van Bruinessen (1990) shows that the corpus itself is historically structured and selective. Transmission is therefore treated here as the movement and re-performance of textual knowledge through material texts, teachers, students, scripts, and institutional routines. This formulation avoids reducing textual transmission to the persistence of a method while still allowing pedagogical practices to be analysed as mechanisms through which access to the corpus is organised.

2. Sorogan, Bandongan/Wetonan, and Grammatical Support for Arabic Reading

Sorogan generally describes an individual or small-scale encounter in which a student reads or presents a text to a teacher and receives correction. *Bandongan* or *wetonan* generally describes a collective session in which the teacher reads, translates, parses, or comments while students follow and annotate. Kamal's (2020) library study conceptualises these as individual and collective learning modes, while field studies at Tanwirunnida' in Magelang and PDF Minhajurrosyidin in Jakarta document both methods within the same institutional setting (Anshori & Wardana, 2022; Yusuf, Anggara, & Sayuti, 2026). From an Arabic-reading perspective, their potential complementarity lies in different distributions of attention. Collective exposition can model

pronunciation, parsing, lexical choices, and interpretive moves for many learners at once, whereas individual presentation can expose errors that remain invisible in a large-group setting. The studies justify analysing these functions together, but they do not demonstrate that every *pesantren* uses the same pacing, correction criteria, or balance between the two modes. Functional complementarity must therefore be distinguished from a claim of universal institutional design.

Nahwu and *sharaf* enter this repertoire both as bodies of explicit knowledge and as resources mobilised during reading. Mariyam's (2021) quantitative study at *Pesantren Riyadhul Huda* reports a statistically significant association between mastery of *nahwu-sharaf* and *kitab-reading* ability. The finding supports a relationship between grammatical competence and reading performance, but its correlational design does not establish which instructional route produces the relationship or whether stronger readers also acquire grammar more successfully. Qualitative research on *Alfiyah* memorisation at *Ma'had Aly Annur II Al-Murtadlo* describes memorisation, *setoran*, teacher correction, and application of remembered rules in reading, but its own design is qualitative and should not be used as a source of correlation coefficients (Saleh, 2025). Nasution's (2024) case at *Tanjung Pasir Al-Awwabien* similarly documents memorisation and *bandongan* within *nahwu* instruction. Taken together, these studies support a view of grammatical knowledge as distributed across memory, explanation, correction, and application rather than as an isolated prerequisite that automatically generates comprehension. A separate correlational study at *Al Hikmah Tasikmalaya* found that *nazam* memorisation was associated with students' understanding of *qawa'id*, although the design did not test *kitab-reading* outcomes directly (Jamilatussa'diyah, Sanah, & Rifa'i, 2022). Broader Arabic literacy research, conducted with Arabic-speaking

readers rather than Indonesian learners, similarly indicates that morphological awareness contributes to reading comprehension and that its development is shaped by the structure and linguistic distance of Arabic varieties (Vaknin-Nusbaum & Saiegh-Haddad, 2020; Shahbari-Kassem, Schiff, & Saiegh-Haddad, 2025). At the level of intellectual history, Yaacob (2025) argues in *Senarai* that Arabic grammar is closely related to textual integrity and the continuity of Islamic knowledge; in the present review, that proposition is used as contextual framing rather than as direct evidence of learning effects.

3. *Makna Gandul*, *Pegon*, and Vernacular Scaffolding in Arabic Reading

Makna gandul is treated in this article as an interlinear or closely attached vernacular glossing practice through which Arabic words, syntactic relations, or interpretive cues are rendered in a local language. *Pegon* refers to the Arabic-derived script commonly used for Javanese and, in varying conventions, other local-language writing. The two intersect, but the distinction is essential: a script is not the same analytical object as a glossing procedure. Regional languages, orthographies, reading orders, grammatical markers, and conventions for connecting Arabic words to vernacular equivalents may differ. For Arabic reading pedagogy, the important question is therefore not whether *Pegon* or *makna gandul* is present in the abstract, but what cognitive and textual work the local convention performs. It may externalise lexical meaning, make syntactic relations visible, preserve an authorised translation, or provide a record that learners can revisit after oral exposition. Each of these functions can support reading, but none should be presumed to operate identically across regions or institutional lineages.

Rahmawati and Negara's (2021) training report at Pondok *Pesantren* Darul Huda Putri identifies unfamiliarity

with Arab *Pegon* as an obstacle for new students and reports that structured training supported participation in *kitab-kuning* learning. The design is useful for identifying a local access requirement, but it does not establish *Pegon* as a universal prerequisite across Indonesia or isolate its effect from other aspects of the intervention. Jamaluddin's (2019) historical account of Javanese *makna-gandul* books and Ricci's (2011) wider study of Arabic cosmopolitan textual circulation provide a broader interpretive frame. They make it possible to understand vernacular glossing not simply as compensation for inadequate Arabic, but as a mode of textual participation in which Arabic and local languages are materially and intellectually connected. Tuncer's (2024) comparative discussion further illustrates variation across Malay and Javanese settings. For this review, vernacular scaffolding is therefore treated as a locally situated reading technology whose contribution to comprehension and textual transmission must be established context by context.

4. Variation and Alternative Pathways to *Kitab-Kuning* Reading Competence

A defensible account must include cases that do not reproduce the anticipated configuration. Priyatna, ZA, and Barni (2024) describe video, e-kitab, and online forums as supports layered onto *sorogan* and *bandongan*, while also reporting infrastructure, cultural-resistance, and digital-skill constraints. The case is important because it shows that inherited reading routines can be supplemented without being mechanically replaced. Sholihin (2024) examines the Nashar method at *Assalafi Al Fithrah* Surabaya as a simplified, Indonesian-mediated and song-supported approach to *nahwu-sharaf*. Habibillah and Ulum (2025) report a mixed-method evaluation of *Amtsilati* at Darun Najah Jember and present quantitative learning gains alongside qualitative evidence. These alternatives

demonstrate that access to kitab reading can be reorganised around new mnemonic sequences, simplified grammatical presentation, or digitally mediated support. They therefore function as negative cases against any model that treats classical *matn* memorisation or a fixed *sorogan-bandongan* sequence as a necessary condition of Arabic reading competence.

Variation is analytically productive because it shifts the central question from preservation of methods to configuration of functions. If traditional practices can coexist with digital media, and if alternative grammar methods can be introduced to accelerate access to kitab reading, then continuity is better conceptualised as selective reproduction under changing constraints. Some elements may be retained because they serve verification or authority functions; others may be modified because learners need faster grammatical access, teachers need scalable materials, or institutions gain new media resources. The empirical question becomes which components are retained, reconfigured, supplemented, or displaced in a particular institution and with what consequences for reading. This also protects the analysis from assuming that innovation necessarily moves pedagogy away from textual tradition. In several reviewed cases, the valued outcome remains engagement with *kitab-kuning* texts even when the instructional mechanism changes. That distinction becomes central to the transmission-ecology model developed below.

Theoretical Framework: Cultural Reproduction in Arabic Reading Pedagogy

This article uses Bourdieu and Passeron's (1990) theory of cultural reproduction as its principal interpretive frame. It is not employed as a theory of second-language reading acquisition and does not explain, by itself, the cognitive processes through which learners decode or comprehend Arabic. Its role is different: it helps explain

how institutions reproduce valued forms of textual competence, authority, and legitimate scholarly practice across generations. Importantly, reproduction is not equated with conscious preservation. Educational arrangements can reproduce recognised forms of knowledge through routine classification, embodied dispositions, assessment, and unequal access to cultural capital even when no actor has formulated an explicit preservation policy. This distinction is particularly useful for the present review because the evidence often documents recurrent reading practices and criteria of competence but rarely provides direct institutional statements explaining why those practices are retained.

Four observable implications guide the synthesis. First, if *kitab-kuning* literacy constitutes valued cultural capital, studies should show that competence in reading, parsing, translating, or explaining texts has consequences for recognised achievement or scholarly standing. Second, reproduction should be visible in patterned relations between authorised teachers, prescribed or recurrent texts, correction routines, and criteria of competent reading. Third, the distribution of that competence should depend on resources such as time, teacher expertise, script familiarity, access to particular texts, and opportunities for individual feedback. Fourth, reproduction should remain compatible with change: digital media or alternative grammar methods may alter the mechanism while leaving the valued object of competence substantially intact. These implications allow cultural reproduction to be assessed against the reviewed evidence rather than used as a decorative label. They also keep the analysis tied to Arabic reading pedagogy by asking how institutional arrangements shape opportunities to acquire and demonstrate textual competence.

The framework is tested against rival explanations rather than used as an endorsement of continuity.

Institutional path dependence predicts persistence because routines, schedules, texts, and teacher competencies are costly to replace; resource constraints predict reliance on scalable collective methods; teacher formation predicts reproduction because instructors may teach through the methods in which they themselves acquired competence; assessment traditions predict persistence where progression depends on parsing, translation, or kitab reading; and explicit heritage commitment predicts direct actor statements that particular methods are retained to preserve a lineage or identity. Only the last explanation warrants a strong claim of deliberate preservation, and the reviewed evidence rarely supplies that kind of direct statement. The discussion therefore distinguishes observed function, plausible mechanism, and demonstrated intention. This distinction also leaves room for pedagogical problem-solving: when Nashar, Amsilati, or digital resources are adopted, adaptation may reflect learner difficulty, time constraints, institutional experimentation, or changes in media rather than a simple opposition between tradition and modernity.

Method

1. Review Design and Scope

This study is a critical integrative review of heterogeneous empirical, historical, and conceptual literature. It does not claim exhaustive systematic coverage and is not presented as a PRISMA review or historical genealogy. The design follows the basic logic of integrative review methodology, which permits diverse empirical and theoretical sources to be compared while requiring explicit attention to problem identification, source discovery, data evaluation, analysis, and presentation (Whittemore & Knafl, 2005). In this article, that logic is adapted to a humanities-and-education problem in which qualitative case studies, quantitative associations, mixed-method evaluations,

library research, training reports, and material-text scholarship answer different kinds of questions. The purpose is not to estimate a pooled effect but to determine what each design can legitimately support about Arabic reading pedagogy and textual transmission.

The integrative design is appropriate because the research question spans instructional practice, linguistic competence, textual materiality, vernacular mediation, and institutional continuity. These evidence types are not interchangeable units. A quantitative association can establish a measured relationship but may say little about classroom mechanism; a qualitative case can describe sequences of correction and application but cannot provide a population-level effect; a historical study can clarify the material and regional development of glossed texts without proving how contemporary learners process them; and a conceptual article can sharpen definitions without establishing prevalence. The review therefore privileges claim-to-design fit over citation frequency. This design choice directly addresses the risk of treating all sources as equivalent votes and allows contradictory or alternative cases to constrain the synthesis.

2. Source Discovery, Eligibility, and Corpus Architecture

Source discovery and verification were conducted during July-August 2026 across Google Scholar, Garuda, Moraref, institutional journal or repository platforms, publisher pages, and citation-oriented searches around foundational and directly relevant works. Search terms included combinations and variants of *sorogan*, *bandongan*, *wetonan*, *kitab kuning*, *kitab gundul*, *nahwu sharaf*, *qawa'id*, *ilmu alat*, *hafalan matan*, *Alfiyah*, *Ajurumiyyah*, *makna gundul*, *makna pesantren*, *Pegon*, Arabic reading, and *kitab-reading* competence. Google Scholar was treated as a discovery environment rather than a source from which a stable numerical yield could be claimed. Backward and forward

citation-oriented searches around van Bruinessen (1990), Ricci (2011), and directly relevant pedagogical studies were used to identify historical, material-text, and practice-specific works that could discipline terminology or challenge the emerging model.

A source was eligible for the practice-specific evidence set when it (a) had identifiable authorship and stable bibliographic metadata; (b) examined an Indonesian *pesantren*, *salaf*/traditional Islamic educational setting, or a directly relevant *kitab-reading* practice; (c) provided enough methodological information to distinguish its design; and (d) supplied a finding directly relevant to at least one analytic relation in this review. For reading-focused synthesis, the extraction also required that the source permit identification of at least one relevant dimension of Arabic reading competence where reported, such as vocalisation, grammatical parsing, translation, comprehension, interpretation, or a locally defined composite of *kitab-reading* ability. Historical, bibliographic, conceptual, disability/access, and cross-regional works were not allowed to function as practice-specific empirical evidence unless they met the same conditions. They were retained, where useful, in a separate contextual corpus.

During source-level verification, each practice-specific record was checked against publisher or full-text information for authorship, outlet metadata, design, institutional context, and the exact finding used in the synthesis. Records addressing general Arabic education, medieval Qur'anic memorisation, or broad *pesantren* themes without direct engagement with the focal practices were not used as practice-specific evidence. General and comparative works were retained only for contextual interpretation. Empirical results were attributed to the original study rather than to secondary citations wherever direct verification was possible; the same verification procedure governed Table 1 and the bibliography. This separation is

important because a general historical or theoretical source may explain why a category matters without supplying evidence that a particular instructional practice produces a particular reading outcome.

3. Evidence Weighting and Synthesis

Instead of a single quality score, evidence was weighted according to design fit and claim directness. Quantitative or mixed-method studies were used for measured associations or reported learning change, with causal language avoided when designs did not support it. Qualitative field studies were used to establish observed practices, actor descriptions, implementation conditions, local sequencing, and institutional combinations. Library research and conceptual articles were used for definitions and interpretation, not for empirical prevalence. Training or community-service reports were treated as local implementation evidence with limited inferential strength. Historical and material-text scholarship was used to discipline terminology and contextualise texts, scripts, and vernacularisation rather than to infer present-day pedagogy. Within each category, stronger weight was given to claims supported directly by the study's own data and design rather than to statements repeated from secondary literature.

Synthesis proceeded in two stages. First, descriptive charting recorded setting, design, focal practice, textual object where reported, dimension of Arabic reading competence addressed, outcome, and design-specific limitation. Reading dimensions included accurate or vocalised reading, grammatical parsing, translation, comprehension, interpretation/explanation, and locally defined *kitab-reading* competence when a study used a composite measure. Second, an abductive synthesis compared patterns across studies and actively searched for disconfirming or alternative configurations. The higher-

order themes reported below—paired verification and exposition, distributed grammatical support, vernacular scaffolding, and adaptive plurality—are therefore not identical to the search terms. A proposed relation was retained only when it could be linked to study-level evidence and bounded by contrary or weaker cases. The final transmission-ecology model is thus an analytical synthesis of recurrent relations, not a claim that the reviewed sources collectively document one historical or causal system.

4. Audited Evidence Base

Table 1 separates direct practice evidence from contextual material and makes visible the design, setting, evidentiary use, and principal limitation of each study. This structure prevents a general or comparative source from generating an artificial gap merely because it does not address a preselected practice. It also shows where the evidence speaks directly to Arabic reading outcomes and where it only clarifies implementation or context. Table 2 then defines the key terms used in the synthesis so that differences between *kitab kuning* and *kitab gundul*, *makna gundul* and *Pegon*, or memorisation and comprehension are not collapsed. Together, the two tables form the record-level audit trail for the claims developed in the Findings section.

Table 1. Characteristics and Evidentiary Weight of Studies on *Kitab-Kuning* Arabic Reading Practices in Indonesian *Pesantren*

Source	Setting / design	Direct evidentiary use	Principal limitation / weight
Kamal (2020)	Library research; <i>pesantren</i> tradition	Conceptualises <i>sorogan</i> as individual and <i>bandongan</i> as collective learning modes	No field sample; contextual, not prevalence evidence

Source	Setting / design	Direct evidentiary use	Principal limitation / weight
Mariyam (2021)	<i>Pesantren</i> Riyadhul Huda; quantitative, total sampling and tests	Association between <i>nahwu-sharaf</i> mastery and <i>kitab-reading</i> ability	Correlational; does not identify instructional causality
Rahmawati & Negara (2021)	Darul Huda Putri; <i>Pegon</i> training / community intervention	<i>Pegon</i> unfamiliarity as a local barrier; training supports participation in <i>kitab</i> learning	Local training report; limited outcome controls
Anshori & Wardana (2022)	Tanwirunnida, Magelang; descriptive qualitative case study	<i>Bandongan</i> and <i>sorogan</i> used together in planning, implementation, and evaluation of <i>kitab</i> reading	Single institutional case; no comparison group
Nasution (2024)	Tanjung Pasir Al-Awwabien, Jambi; qualitative descriptive	Memorisation and <i>bandongan</i> combined in <i>nahwu</i> instruction; reports supporting and inhibiting factors	Single case; self/actor-reported instructional outcomes
Priyatna, ZA, & Barni (2024)	<i>Salafiyah pesantren</i> ; qualitative case study with interviews and observation	Digital media can supplement <i>sorogan/bandongan</i> ; infrastructure, resistance, and digital skill constrain adoption	Context-specific; proposed optimisation exceeds causal evidence
Sholihin (2024)	<i>Assalafi</i> Al Fithrah, Surabaya; qualitative descriptive	Nashar method provides an alternative simplified/song-supported <i>nahwu-sharaf</i> pathway linked to <i>kitab</i> reading	Single case; effectiveness primarily qualitative
Saleh (2025)	<i>Ma'had Aly</i> Annur II Al-Murtadlo, Malang; qualitative field study	<i>Alfiyah</i> memorisation, <i>setoran</i> , correction, and application in <i>kitab</i> reading are institutionally connected	Qualitative; statistical coefficient in prior version was not this study's result
Habibillah & Ulum (2025)	Darun Najah, Jember; concurrent-embedded mixed method	Amtsilati as an alternative grammar pathway; reports post-test mean 82.40 and N-Gain 0.95	Local evaluation; generalisability and causal attribution remain bounded by design
Yusuf, Anggara, & Sayuti (2026)	PDF Minhajurrosyidin, Jakarta; descriptive qualitative with observation, interviews, documents	Routine <i>bandongan/sorogan</i> implementation; teacher training and evaluation standards still developing	Formal <i>diniyah</i> setting, not interchangeable with all <i>salaf pesantren</i>

Table 2 clarifies the operational distinctions used throughout the synthesis. These definitions are deliberately narrower than common shorthand and should be read as analytical conventions rather than claims that every *pesantren* uses the same terminology.

Table 2. Operational Definitions of Core Terms in *Kitab-Kuning* Arabic Reading Pedagogy

Term	Operational use in this article	What is not assumed
<i>Kitab kuning</i>	<i>Pesantren</i> -associated corpus of Islamic scholarly and teaching books in Arabic script, materially and disciplinarily diverse	Not synonymous with yellow paper, unvocalised text, or <i>matn-plus-sharh</i> format
<i>Kitab gundul</i>	An Arabic text/edition substantially lacking vowel marks	Not a synonym for the whole <i>kitab-kuning</i> corpus
<i>Sorogan</i>	Individual or small-scale student presentation/reading subject to teacher correction	Not necessarily identical in pacing, assessment, or group size across institutions
<i>Bandongan / wetonan</i>	Collective teacher-led reading, translation, parsing, or commentary followed by students	Not presumed pedagogically passive in every setting or always paired with <i>sorogan</i>
<i>Hafalan matan</i>	Memorisation of a concise teaching text or versified grammatical material such as <i>Alfiyah</i>	Not treated as equivalent to Qur'anic memorisation or as proof of comprehension
<i>Makna gandul</i>	Interlinear or closely attached vernacular glossing/reading practice	Not synonymous with <i>Pegon</i> and not uniform across regions
<i>Pegon</i>	Arabic-derived script used for local-language writing in Indonesian contexts	Not treated as identical with Jawi or as nationally standardised
<i>Kitab-reading competence</i>	Context-specific capacity to vocalise, parse, translate, explain, or otherwise read designated kitab according to local criteria	Not reduced to one nationally standard outcome instrument

Findings

1. The Evidence Base Is Local and Methodologically Uneven

The first finding is methodological rather than substantive: the available evidence does not justify a national claim about a single *pesantren* system. The practice-specific set is dominated by single-site qualitative studies,

supplemented by one quantitative association study, a mixed-method local evaluation, library research, and a training report. Most sites are located on Java, with additional evidence from Jambi and other Indonesian settings; institutional types vary from *salaf*-oriented *pesantren* to *Ma'had Aly* and Pendidikan Diniyah Formal. The textual objects and outcome measures also differ. Some studies focus on *nahwu* instruction, some on *kitab-reading* performance, some on *Pegon* access, and others on the organisation of *sorogan* and *bandongan*. The synthesis must therefore use formulations such as 'in the reviewed cases' or 'a recurrent configuration' rather than 'the *pesantren* tradition' as if it were homogeneous. This limitation is not merely geographical; it concerns the comparability of what counts as reading competence and instructional success.

Sources also carry different evidentiary weight for Arabic reading questions. Mariyam (2021) can support a measured relationship between grammatical mastery and reading competence but cannot establish which teaching method caused it. Saleh (2025) can show how *Alfiyah* memorisation, *setoran*, correction, and reading were locally connected but cannot supply a population-level effect size. Rahmawati and Negara (2021) can show that *Pegon* unfamiliarity impeded new students in one female *pesantren* and that training was locally useful, but it cannot establish a universal literacy sequence. Anshori and Wardana (2022) and Yusuf, Anggara, and Sayuti (2026) can document co-presence and implementation of *sorogan* and *bandongan* but not historical origin or causal superiority. Treating these designs as equivalent would obscure their design-specific limits. The more defensible strategy is to use each study for the claim it can actually support and then examine whether relations recur across independently bounded contexts.

2. *Sorogan* and *Bandongan* as Complementary Modes of Arabic Reading Instruction

The most consistent cross-study pattern concerns *sorogan* and *bandongan*. Kamal (2020) describes their conceptual distinction as individual versus collective learning service. Anshori and Wardana (2022) and Yusuf, Anggara, and Sayuti (2026) document both within the same institution, while the latter study also shows that routine implementation can coexist with developing teacher-training and evaluation systems. Together these studies support functional complementarity in Arabic reading instruction: one mode can provide scalable exposition of pronunciation, parsing, translation, and commentary, while the other creates opportunities for more individualised verification and correction. The asymmetry matters because the two modes distribute teacher attention differently. *Bandongan* can expose many learners to an authorised reading at once, whereas *sorogan* can make individual errors visible. What the evidence does not establish is that the pair was historically designed as one integrated system or that every *pesantren* deploys both in the same balance.

Priyatna, ZA, and Barni (2024) adds an important negative case to a static interpretation. Digital video, e-kitab, and online forums can be layered onto traditional modes while technical infrastructure, cultural resistance, and digital competence shape what is feasible. This case supports adaptation within continuity, but it also demonstrates that the instructional mechanism is contingent on material resources. From an Arabic-reading perspective, digital supplementation potentially changes where modelling, repetition, annotation, or access to text occurs without necessarily changing the valued outcome of competent engagement with kitab. The implication is that *sorogan* and *bandongan* should not be treated as fixed technologies. Their functions—verification, exposition, modelling, and guided interpretation—may be redistributed across face-to-face and digital media. Persistence is therefore compatible with change and cannot

be reduced either to stagnation or to deliberate heritage policy.

3. Grammatical Knowledge, Memory, and Application in Arabic Reading Development

The reviewed evidence does not support a single linear pathway from memorisation to independent reading. It supports a looser distribution of grammatical work across memory, explanation, correction, and application. Mariyam (2021) finds that higher *nahwu-sharaf* mastery is associated with stronger *kitab-reading* ability. Nasution (2024) shows memorisation and *bandongan* operating together in a local *nahwu* curriculum. Saleh (2025) documents repeated memorisation of *Alfiyah*, direct *setoran* to an *ustadz*, correction of errors, and subsequent application of remembered material during *kitab* reading. Taken together, these sources make it plausible that grammatical memory becomes pedagogically useful when it can be mobilised in reading, especially when learners receive explanation and corrective feedback. They do not show that memorisation alone produces comprehension, nor do they establish that every dimension of reading competence depends on the same amount or type of explicit grammatical knowledge.

Alternative methods prevent the synthesis from mistaking one route for the only route. Sholihin (2024) presents Nashar as a simplified and mnemonic method for beginners, while Habibillah and Ulum (2025) evaluate Amsilati as a structured grammar approach and report substantial learning gains in one *pesantren*. These studies complicate any claim that classical *matn* memorisation is a necessary condition of *kitab* literacy. The more defensible higher-order theme is distributed grammatical support: different institutions externalise, memorise, explain, practise, or accelerate grammar through different combinations of texts and methods. For Arabic reading pedagogy, this means grammar should be analysed by

function rather than by prestige of form. The critical question is how a method enables learners to recognise syntactic relations, assign appropriate endings, resolve ambiguity, connect lexical meaning to sentence structure, and apply those resources during actual reading. The current evidence suggests multiple pathways, but comparative designs are still needed to determine how those pathways differ in learning demands and outcomes.

4. Vernacular Glossing as Scaffolding for Arabic Text Comprehension

Evidence for *makna gandul* and *Pegon* is qualitatively different from evidence for *sorogan-bandongan*. Rahmawati and Negara (2021) directly address *Pegon* training for new students, while Jamaluddin (2019) documents a historical publishing culture of Javanese *makna-gandul* books. These sources show that vernacular inscription can mediate access to Arabic texts and can become a durable textual form in its own right. In pedagogical terms, an interlinear gloss may reduce the immediate burden of lexical retrieval, preserve a teacher-mediated interpretation, and make some grammatical relations visible through locally conventional markers. Yet the reviewed literature does not provide experimental evidence isolating these functions or demonstrating how long learners depend on them. The evidence therefore supports scaffolding as a plausible and locally documented function, not a universal developmental stage from dependence to independent Arabic reading.

Ricci's (2011) account of translation in the Arabic cosmopolis helps interpret this practice without treating it as a temporary deficiency. Vernacularisation can be a mode of participation in an Arabic-oriented intellectual world rather than an abandonment of it. Tuncer (2024) similarly indicates variation across Malay and Javanese examples, while Jamaluddin (2019) shows that *makna-gandul* formats

can be materially stabilised through publication. Consequently, *makna gandul* is best analysed as a locally situated technology of interpretation and inscription whose relationship to grammar, teacher authority, and independent reading must be demonstrated case by case. This finding is important for Arabic language education because it broadens the concept of reading support beyond dictionaries or explicit grammar explanation. It also matters for textual transmission because the vernacular gloss itself becomes part of the material pathway through which a text is learned, remembered, and re-read in a local scholarly environment.

5. A Transmission-Ecology Model of Arabic Reading Pedagogy

Across the audited evidence, four relational components recur sufficiently to justify a bounded transmission-ecology model of Arabic reading pedagogy. The first is a textual object or corpus that defines what counts as advanced or legitimate reading. The second is grammatical support, supplied through explicit *nahwu-sharaf* study, memorised *matn*, simplified methods, or combinations of these. The third is teacher-mediated verification and exposition, especially visible in *sorogan* and *bandongan*. The fourth is mediation through local scripts, glosses, or newer digital media. *Kitab-reading* competence functions as an important outcome, but its dimensions and assessment vary by institution. The model therefore treats reading as an activity distributed across learner knowledge, material texts, teacher guidance, annotation systems, and institutional expectations rather than as a purely individual linguistic skill detached from its learning environment.

The model is intentionally bounded. No reviewed study demonstrates all components as one historically coordinated mechanism across Indonesian *pesantren*. Nasution (2024) provides direct evidence for more than one

practice within a single grammar curriculum; the *sorogan-bandingan* cases show another local combination; and the *Pegon* literature shows a distinct layer of vernacular mediation. Alternative grammar methods and digital supplementation further demonstrate that components can be substituted or reconfigured. The ecology is therefore an abductive model that integrates recurrent relations across studies, not a claim of national uniformity. Its value is explanatory and heuristic: it identifies where future multi-site research should look for relations among textual material, linguistic resources, teacher authority, scaffolding, media, and reading outcomes, while also allowing cases in which one or more components are absent to challenge the model.

Discussion

1. *Kitab-Kuning* Literacy as a Distinctive Form of Arabic Reading Pedagogy

The central contribution of the synthesis is to reposition *kitab-kuning* instruction as a distinctive form of Arabic reading pedagogy rather than to treat it only as a marker of *pesantren* tradition. Across the reviewed cases, reading is not reduced to decoding written words. It can require appropriate vocalisation, grammatical parsing, lexical and contextual translation, explanation, and alignment with an authorised interpretation. Those demands help explain why explicit *nahwu-sharaf* knowledge, teacher-mediated correction, collective modelling, memorised rules, and vernacular annotation can remain pedagogically salient. Each practice distributes part of the reading task across different resources: grammar supplies analytic categories, *sorogan* intensifies individual verification, *bandongan* models interpretation at scale, and *makna gandul* can externalise information that learners have not yet fully internalised. Seen in this way, the practices

become analytically comparable because they contribute to different dimensions of the same broad reading problem.

This reading-oriented interpretation also prevents textual transmission from becoming an abstract heritage claim. Textual continuity depends on readers who can enter, parse, translate, and explain the texts that institutions value. At the same time, the form in which competence is cultivated is shaped by the material text, local annotation conventions, and criteria of acceptable reading. The transmission-ecology model therefore connects two domains that are often separated: Arabic language education and Islamic textual scholarship. It does not claim that pedagogical effectiveness and heritage continuity are identical outcomes. Rather, it proposes that the continued circulation of *kitab-kuning* texts is partly mediated by pedagogical arrangements that make those texts learnable and assessable within particular institutions. This connection is strongest where the evidence directly documents reading practices and weakest where historical motives or long-term outcomes are inferred rather than observed.

2. Cultural Reproduction and Rival Explanations

The theoretical distinction that remains essential is between continuity-producing functions and demonstrated intention. The reviewed literature supports the former more strongly than the latter. *Sorogan* can reproduce a relation of individual verification; *bandongan* can reproduce authoritative collective exposition; memorised grammatical texts can reproduce specialised linguistic resources; and vernacular glossing can reproduce locally intelligible ways of entering Arabic texts. These are effects and functions. Bourdieu and Passeron (1990) make this distinction useful because reproduction can occur through embodied dispositions, routine evaluation, institutional recognition, and differential possession of cultural capital without

transparent institutional deliberation. The presence of a continuity-producing function therefore does not establish that an institution consciously selected a method in order to preserve heritage.

Several rival mechanisms fit the available evidence. Resource constraints are visible where collective instruction scales more easily than individual correction and where digital adaptation is limited by infrastructure or skills (Priyatna, ZA, & Barni, 2024). Teacher formation is plausible because the competence required to teach and assess kitab reading is itself acquired through institutional routines. Assessment traditions are relevant where the ability to parse, translate, or explain texts becomes a criterion of progression. Path dependence is plausible because replacing texts, teacher competencies, schedules, and evaluation criteria simultaneously is organisationally costly. Yusuf, Anggara, and Sayuti (2026) further show that implementation quality depends on standardisation, training, and evaluation rather than simply on loyalty to tradition. Explicit preservation remains a possible explanation, but it requires direct evidence such as curricular rationales, policy statements, teacher testimony, prefaces, or archival material that the present corpus rarely provides.

3. Multiple Pathways to Arabic Reading Competence

The evidence does not support a total opposition between traditional and communicative or technology-mediated education. The focal practices are oriented primarily toward reading, parsing, interpretation, and authorised engagement with classical texts; communicative approaches optimise partly different language outcomes. These objectives can compete for curricular time, but they can also coexist. Priyatna, ZA, and Barni (2024) directly document digital supplementation rather than simple replacement, while the alternative-method studies show

internal pedagogical innovation within *kitab-reading* goals. This means that the relevant analytical contrast is not tradition versus modernity but instructional configuration versus specified learning outcome. A method should be examined in relation to what dimension of Arabic competence it is designed to support and under what institutional conditions it is implemented.

Nashar and Amtsilati sharpen this conclusion because they demonstrate that grammatical support can be reorganised when institutions seek easier or faster access to reading. If a method changes while the valued outcome of kitab reading remains, cultural reproduction may operate at the level of competence and textual orientation rather than at the level of a fixed teaching routine. The same logic applies to digital media: video or e-kitab may alter opportunities for repetition, access, or annotation without displacing teacher authority or the status of the text. Future comparison should therefore distinguish at least three questions: which reading dimensions are targeted, which pedagogical mechanisms are used to support them, and which textual or institutional values remain stable when the mechanism changes. Such a framework can accommodate innovation without assuming either decline of tradition or automatic pedagogical progress.

4. Authority, Access, and Uneven Learning Conditions

A reproduction perspective also prevents continuity from being treated normatively as an unqualified good. Individual correction requires teacher time; collective sessions can reduce opportunities for immediate participation; script familiarity can differentiate newcomers; explicit grammar can impose substantial memory and analytic demands; and digital innovations depend on infrastructure and skills. These are not external concessions to an otherwise coherent model but conditions that shape who can acquire the valued literacy and how quickly.

Teacher authority can provide reliable correction and interpretive guidance, yet it can also concentrate control over pacing, acceptable translation, and criteria of competence. The reviewed evidence is not sufficient to determine whether those features improve or constrain learning across populations, but it is sufficient to show that access to reading practice is institutionally organised rather than evenly available by default.

The available evidence is also insufficient to generalise inequalities across gender, class, disability, and geography. Accordingly, the discussion is limited to directly documented issues: variation in learner background, teacher training, evaluation standards, infrastructure, digital skill, motivation, and prior *Pegon* familiarity. This restraint is methodologically important because a transmission model can easily become romantic if it describes continuity without asking who has access to the time, texts, teachers, and script knowledge required to participate. Future research should therefore examine distributional questions directly rather than infer them from unrelated evidence. In Arabic reading research, that means reporting learner entry competence, institutional resources, instructional intensity, feedback opportunities, and assessment criteria alongside the method being studied. Such data would make it possible to distinguish problems inherent to a pedagogical mechanism from problems created by weak or unequal implementation.

5. Implications for Arabic Reading and Textual-Transmission Research

The transmission-ecology model generates a field-research agenda rather than a national policy prescription. A multi-site comparative design could record the specific kitab and edition studied; degree of vocalisation; language and script of annotation; teacher and student roles; use of *sorogan*, *bandongan/wetonan*, memorisation, parsing, and

glossing; dimensions of Arabic reading competence assessed; digital supplements; teacher formation; and explicit statements of institutional rationale. Such data would permit researchers to test whether the same pedagogical function is realised through different methods and whether different configurations produce different profiles of reading competence. They would also allow cultural reproduction to be distinguished from explicit preservation, resource constraint, teacher formation, assessment tradition, or path dependence rather than inferred from persistence.

Historical claims require a different but complementary evidentiary design. Researchers wishing to retain the language of genealogy should trace bounded texts, editions, marginalia, curricula, teachers, *ijazah* or *sanad* materials, institutional archives, and changing technologies across a defined period. For Arabic language education, the most productive intersection may be to examine how changes in material text and instructional mediation alter the skills required of readers. For Islamic textual scholarship, the corresponding question is how pedagogical arrangements shape the circulation and local interpretation of a corpus. The present review does not perform that historical reconstruction. Its contribution is to formulate a set of pedagogical and textual relations that can be tested in future fieldwork and historical research, thereby connecting *mahārah al-qirā'ah* research with the study of knowledge transmission in Indonesian *pesantren* without collapsing one field into the other.

Conclusion

This integrative review shows that Arabic reading competence in Indonesian *pesantren* is supported by combinations of pedagogical practices rather than by a single instructional method. Across the reviewed studies, *sorogan* and *bandongan/wetonan* recurrently function as

complementary modes of teacher-mediated reading instruction, with *sorogan* providing opportunities for individual verification and correction and *bandongan* facilitating collective exposition of Arabic texts. Mastery of *nahwu-sharaf* provides an important grammatical resource for kitab reading, while memorisation of grammatical *matn*, *makna gandul*, and Pegon-based glossing can provide additional forms of support. The evidence does not, however, support one linear pathway from grammatical memorisation to reading competence, nor does it establish that all *pesantren* employ the same configuration. Alternative grammar methods and digital supplementation demonstrate that the mechanisms of Arabic reading pedagogy can change while engagement with valued Arabic scholarly texts remains a central outcome in the documented settings.

The study therefore proposes transmission ecology as a bounded and falsifiable analytical model rather than as a historical fact or institutional intention. Its components—textual object, grammatical resources, teacher-mediated verification and exposition, vernacular scaffolding, and adaptive media—can be present in different combinations and with different consequences for vocalisation, parsing, translation, comprehension, and interpretation. Cultural reproduction provides a way to explain how valued forms of textual competence and authority may persist without assuming conscious preservation as the sole cause. For Arabic language education, the principal contribution is to frame *kitab-kuning* instruction as a distinctive form of Arabic reading pedagogy in which linguistic knowledge, guided textual practice, teacher correction, vernacular mediation, and institutional expectations interact. For Islamic textual studies, the model highlights how those pedagogical arrangements can sustain and reconfigure locally embedded pathways of textual transmission.

The conclusions remain bounded by a heterogeneous and predominantly local evidence base. The review is integrative rather than exhaustive, and most included studies cannot identify causal mechanisms, compare instructional configurations directly, or demonstrate institutional intention. Future research should therefore combine multi-site fieldwork with clearer Arabic reading assessment frameworks that distinguish accurate reading, grammatical parsing, translation, comprehension, and interpretation. Historical work should separately trace texts, editions, glosses, teachers, scripts, curricula, and institutional change when claims of genealogy or long-term continuity are made. The most productive next step is not to ask whether *pesantren* reading methods are simply traditional or modern, but to examine how particular combinations of pedagogical practice develop Arabic reading competence while simultaneously sustaining, modifying, or reconfiguring the textual traditions associated with *kitab-kuning* learning.

Bibliography

- Aliyah, A. (2018). Pesantren tradisional sebagai basis pembelajaran nahwu dan sharaf dengan menggunakan kitab kuning. *Al-Ta'rib: Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya*, 6(1), 1–25.
<https://doi.org/10.23971/altarib.v6i1.966>
- Anshori, M., & Wardana, B. E. (2022). Implementasi metode bandongan dan metode sorogan dalam pembelajaran kitab kuning di Pondok Pesantren Tanwirunnida', Kecamatan Mungkid, Kabupaten Magelang. *Seminar Nasional Paedagoria*, 2, 292–302.
- Bourdieu, P., & Passeron, J. C. (1990). *Reproduction in education, society and culture* (2nd ed., R. Nice, Trans.). Sage. (Original work published 1970)

- Dhofier, Z. (2011). Tradisi pesantren: Studi pandangan hidup kyai dan visinya mengenai masa depan Indonesia. LP3ES.
- Habibillah, M. J. S., & Ulum, M. B. (2025). Efektivitas metode Amsilati terhadap keterampilan membaca kitab kuning di Pondok Pesantren Darun Najah Umbulsari Jember. *IQRO: Journal of Islamic Education*, 8(2), 762–776. <https://doi.org/10.24256/iqro.v8i2.7681>
- Jamaluddin. (2019). Kitab makna gandul Jawan di Jawa paruh kedua abad ke-XX: Melestarikan tradisi dan meneguhkan ideologi. *Prosiding Mukhtar Pemikiran Santri Nusantara 2018: Revitalisasi Keilmuan Pesantren*. <https://doi.org/10.31219/osf.io/39psb>
- Jaunanto, A., & Mahliatussikah, H. (2020). Characteristic of evaluation in learning Arabic language at Islamic boarding school Anwarul Huda Malang. *Alsuna: Journal of Arabic and English Language*, 3(2), 179–186. <https://doi.org/10.31538/alsuna.v3i2.721>
- Jamilatussa'diyah, J., Sanah, S., & Rifa'i, I. (2022). Taṭbīq Tarīqat Ḥifẓ al-Naẓam fī Ta'lim al-Qawā'id wa-Atharuhū fī Muyūl al-Tullāb wa-Qudratihim 'alā Fahm al-Qawā'id al-'Arabīyah. *Alsina: Journal of Arabic Studies*, 4(1), 45–68. <https://doi.org/10.21580/alsina.4.1.12479>
- Kamal, F. (2020). Model pembelajaran sorogan dan bandongan dalam tradisi pondok pesantren. *Paramurobi: Jurnal Pendidikan Agama Islam*, 3(2), 15–26. <https://doi.org/10.32699/paramurobi.v3i2.1572>
- Mariyam, S. (2021). Hubungan penguasaan nahwu sharaf dengan kemampuan membaca kitab kuning Pesantren Riyadhul Huda. *Tatsqifiy: Jurnal Pendidikan Bahasa Arab*, 2(1), 71–81. <https://doi.org/10.30997/tjpba.v2i1.2828>
- Nasution, N. C. (2024). Penerapan metode menghafal dan bandongan dalam pembelajaran kitab nahwu di Pondok Pesantren Tanjung Pasir Al-Awwabien Jambi.

- Dirasatul Ibtidaiyah*, 4(1), 1–18.
<https://doi.org/10.24952/ibtidaiyah.v4i1.10590>
- Priyatna, S. E., ZA, A. M., & Barni, M. (2024). Menynergikan tradisi dan teknologi: Optimalisasi metode sorogan dan bandongan di pesantren salafiyah melalui media pembelajaran digital. *Bayan lin-Naas: Jurnal Dakwah Islam*, 8(2), 51–71.
<https://doi.org/10.28944/bayanlin-naas.v8i2.1927>
- Rahmawati, I., & Negara, T. D. W. (2021). Pelatihan Arab Pegon bagi santri baru guna meningkatkan kualitas pembelajaran kitab kuning di Pondok Pesantren Darul Huda Putri. *MA'ALIM: Jurnal Pendidikan Islam*, 2(2), 103–112. <https://doi.org/10.21154/maalim.v2i2.3177>
- Ricci, R. (2011). *Islam translated: Literature, conversion, and the Arabic cosmopolis of South and Southeast Asia*. University of Chicago Press.
<https://doi.org/10.7208/chicago/9780226710907>
- Saiegh-Haddad, E., & Schiff, R. (2016). The impact of diglossia on voweled and unvoweled word reading in Arabic: A developmental study from childhood to adolescence. *Scientific Studies of Reading*, 20(4), 311–324.
<https://doi.org/10.1080/10888438.2016.1180526>
- Saleh, A. K. S. (2025). Hafalan Alfiyah Ibnu Malik sebagai strategi pembelajaran nahwu dan penanaman moral mahasantri di pesantren. *Al-Wazan: Journal of Arabic Education*, 3(1), 83–101. <https://doi.org/10.58223/al-wazan.v3i1.416>
- Shahbari-Kassem, A., Schiff, R., & Saiegh-Haddad, E. (2025). Development of morphological awareness in Arabic: The role of morphological system and morphological distance. *Reading and Writing*, 38, 2235–2267. <https://doi.org/10.1007/s11145-024-10581-0>
- Sholihin, C. (2024). Eksplorasi metode Nashar dalam meningkatkan kualitas pembelajaran nahwu sharaf di Pondok Pesantren Assalafi Al Fithrah Surabaya. *EL-BANAT: Jurnal Pemikiran dan Pendidikan Islam*, 14(2),

186–202.

<https://doi.org/10.54180/elbanat.2024.14.2.186-202>

Tuncer, E. (2024). Makna Gandul stilinin Malayca ve Cavaca meal/tefsir çalışmalarında kullanım biçimi ve işlevselliği. *İLSAM Akademi Hakemli Dergisi*, 4(2), 203–230. <https://izlik.org/JA88KK86YN>

Vaknin-Nusbaum, V., & Saiegh-Haddad, E. (2020). The contribution of morphological awareness to reading comprehension in Arabic-speaking second graders. *Reading and Writing*, 33(10), 2413–2436. <https://doi.org/10.1007/s11145-020-10048-y>

van Bruinessen, M. (1990). Kitab kuning: Books in Arabic script used in the pesantren milieu. *Bijdragen tot de Taal-, Land- en Volkenkunde*, 146(2/3), 226–269. <https://doi.org/10.1163/22134379-90003218>

van der Molen, W., Katkova, I. R., & Wieringa, E. (2026). Islamic manuscript culture and transregional Muslim intellectual networks in Mindanao and Southeast Asia. *Senarai: Journal of Islamic Heritage and Civilization*, 2(3), 318–342.

<https://doi.org/10.66236/senarai.2026.2.3.318-342>

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, 52(5), 546–553. <https://doi.org/10.1111/j.1365-2648.2005.03621.x>

Yaacob, S. (2025). Arabic grammar, epistemology, and civilizational thought: A philosophical inquiry into Islamic heritage. *Senarai: Journal of Islamic Heritage and Civilization*, 2(1), 1–13. <https://doi.org/10.66236/senarai.2025.2.1.1-13>

Yusuf, A., Anggara, S., & Sayuti, M. A. (2026). Implementasi metode bandongan dan sorogan dalam pembelajaran kitab kuning di PDF Minhajurrosyidin Jakarta Timur. *Ar Rasyiid: Journal of Islamic Studies*, 4(1), 9–16. <https://doi.org/10.70367/mag3s635>