



## Reconceptualizing Arabic Language Learning in Madrasahs in the Age of Artificial Intelligence

### Merekonseptualisasi Pembelajaran Bahasa Arab di Madrasah di Era Kecerdasan Buatan

**Kartini**

Pendidikan Bahasa Arab, Fakultas Tarbiyah dan Ilmu Keguruan,

IAIN Palopo, Indonesia

e-mail: [kartini@uinpalopo.ac.id](mailto:kartini@uinpalopo.ac.id)

Received: 14-08-2022

Accepted: 25-10-2022

Published: 30-10-2022

#### **How to cite this article:**

Kartini, K. (2022). Reconceptualizing Arabic Language Learning in Madrasahs in the Age of Artificial Intelligence. *Pedagogik Journal of Islamic Elementary School*, Vol. 5(2), 326-347. <https://doi.org/10.24256/pijies.v5i2.4207>

#### **Abstract**

*The rapid diffusion of generative artificial intelligence (AI) is transforming how languages are taught and learned, yet Arabic language instruction in Indonesian madrasahs remains largely organized around grammar memorization, textual decoding, and teacher-centered delivery. This article reconceptualizes Arabic language learning in madrasahs in the age of AI. Employing a library research design, it synthesizes peer-reviewed studies, policy documents, and authoritative books on AI in education, computer-assisted language learning, Arabic pedagogy, Arabic natural language processing, and Islamic educational thought. The data were analyzed through content analysis and thematic analysis. The findings reveal four interrelated shifts: (1) from an instrumental view of Arabic to a human-centered purpose that joins communicative competence, Islamic textual literacy, and AI literacy; (2) from a standard-only, fragmented content model to a layered, usage-oriented one that recognizes diglossia and classical heritage texts; (3) from transmission-based teaching to dialogic, human-AI collaborative pedagogy grounded in scaffolding, interaction, and formative feedback; and (4) from technical adoption to accountable governance that embeds verification (tabayyun), data protection, academic integrity, and equity. The article proposes the HUDA framework (Human-centered, Usage-oriented, Dialogic, Accountable), in which the teacher acts as learning designer and moral exemplar. It concludes that AI can enrich Arabic learning in madrasahs only when its integration is pedagogically purposeful, linguistically critical, and ethically bounded.*

**Keywords:** arabic language learning; madrasah; artificial intelligence; generative ai; islamic education

#### **Abstrak**

*Pesatnya perkembangan kecerdasan artifisial (artificial intelligence/AI) generatif mengubah cara bahasa diajarkan dan dipelajari, sementara pembelajaran bahasa Arab di madrasah Indonesia masih banyak berpusat pada hafalan kaidah, pembacaan teks, dan penyampaian yang berorientasi pada guru. Artikel ini merekonseptualisasi pembelajaran bahasa Arab di madrasah pada era AI. Dengan desain penelitian kepustakaan, artikel ini mensintesis hasil penelitian terpublikasi, dokumen kebijakan, dan buku rujukan tentang AI dalam pendidikan, pembelajaran*

bahasa berbantuan komputer, pedagogi bahasa Arab, pemrosesan bahasa alami bahasa Arab, dan pemikiran pendidikan Islam. Data dianalisis melalui analisis isi dan analisis tematik. Temuan menunjukkan empat pergeseran yang saling terkait: (1) dari pandangan instrumental terhadap bahasa Arab menuju tujuan yang berpusat pada manusia dengan memadukan kompetensi komunikatif, literasi teks keislaman, dan literasi AI; (2) dari model konten yang terfragmentasi dan hanya berfokus pada ragam baku menuju model berlapis dan berorientasi penggunaan yang mengakui diglosia serta teks-teks klasik; (3) dari pengajaran transmisi menuju pedagogi dialogis kolaborasi manusia-AI yang berlandaskan scaffolding, interaksi, dan umpan balik formatif; dan (4) dari adopsi teknis menuju tata kelola akuntabel yang menanamkan verifikasi (*tabayyun*), perlindungan data, integritas akademik, dan keadilan akses. Artikel ini menawarkan kerangka HUDA (Human-centered, Usage-oriented, Dialogic, Accountable) dengan guru sebagai perancang pembelajaran dan teladan moral. Simpulannya, AI dapat memperkaya pembelajaran bahasa Arab di madrasah hanya jika integrasinya bertujuan pedagogis, kritis secara kebahasaan, dan dibatasi oleh etika.

**Kata kunci:** pembelajaran bahasa arab; madrasah; kecerdasan artifisial; ai generatif; pendidikan islam

---

©Pedagogik Journal of Islamic Elementary School. This is an open access article under the [Creative Commons - Attribution-ShareAlike 4.0 International license \(CC BY-SA 4.0\)](#)

## Introduction

Arabic occupies a unique position in Muslim education. As the language of the Qur'an, the Prophetic traditions, and the classical Islamic sciences, it is at once a foreign language for most Indonesian learners and a language of spiritual and intellectual heritage. In Indonesia, madrasahs, the Islamic schools that form part of the national education system under the supervision of the Ministry of Religious Affairs, are the principal formal institutions responsible for teaching Arabic to school-age learners. Arabic is a designated subject within the madrasah curriculum, and successive ministerial decrees have redefined its learning outcomes, most recently under the *Kurikulum Merdeka* (Independent Curriculum) framework adapted for madrasahs.

Despite this policy attention, the literature on Arabic instruction in Indonesia has long described a set of recurring difficulties. Classroom practice has often been characterized by explicit grammar instruction (*nahwu* and *sharaf*), vocabulary lists, and translation exercises, with comparatively little time for listening, speaking, or extended writing (Effendy, 2005). Because learners meet Arabic mainly inside the classroom, opportunities for meaningful input and interaction, which second language acquisition research regards as essential (Krashen, 1982; Long, 1996; Swain, 1985), remain scarce. Recent studies of curriculum implementation show that the movement toward competence-based, learner-centered curricula places new demands on teachers' planning and pedagogical repertoires, while research in Islamic boarding schools

*Pedagogik Journal of Islamic Elementary School*

highlights the continuing effort to integrate the traditional grammar sciences with contemporary approaches (Ardiansyah & Muhammad, 2020). These are not new problems, but they now confront a new technological condition.

Madrasahs are also highly diverse. They include state and private institutions, urban and remote ones, and schools with strong pesantren affiliations alongside those oriented toward general education (Hefner, 2009). Differences in funding, teacher qualifications, and digital infrastructure mean that any proposal to integrate artificial intelligence into Arabic instruction cannot assume uniform conditions.

The new technological condition is the arrival of generative artificial intelligence (AI). Built on transformer architectures (Vaswani et al., 2017) and scaled into large language models (Brown et al., 2020), systems such as ChatGPT can now hold conversations, explain grammatical rules, generate texts at a specified level, add vowel marks, and translate at a speed and cost that would have been unimaginable a decade ago. Education systems worldwide have responded with a mixture of enthusiasm and alarm: enthusiasm over personalized support and reduced teacher workload, and alarm over academic dishonesty, inaccurate output, and learner dependence. Madrasah students, like their peers elsewhere, are likely already using these tools for homework, translation, and text composition, often without guidance. The relevant question for Arabic educators is therefore no longer whether AI will enter the classroom, but on what pedagogical, linguistic, and ethical terms.

Arabic raises particular questions. Its rich derivational morphology, the customary omission of short vowels in writing, and the pervasive diglossia between the standard language and everyday dialects make it a demanding language for natural language processing (Darwish et al., 2021; Farghaly & Shaalan, 2009; Habash, 2010). Generative models are known to produce fluent but false statements, a phenomenon usually called hallucination, and they inherit biases and gaps from their training data (Bender et al., 2021; Weidinger et al., 2021). In a setting where Arabic is the language of revelation, an AI-generated error in a Qur'anic quotation, a hadith attribution, or a grammatical explanation is not a trivial mistake. Educators must therefore weigh the undeniable affordances of AI against distinctive epistemic and moral stakes.

Research on AI in language education has expanded quickly (Yang & Kyun, 2022), and scholarship on Arabic has begun to appear, including studies of AI-supported Arabic teaching and reviews of Arabic chatbots. Three limitations nonetheless stand out.

First, much of the evidence concerns tool performance or short-term learner outcomes rather than the redesign of curricula and pedagogy (Golonka et al., 2014; Zawacki-Richter et al., 2019). Second, relatively few studies situate Arabic learning within the institutional realities and educational philosophy of madrasahs, where language learning is inseparable from religious formation (Al-Attas, 1980; Halstead, 2004). Third, national policy is moving quickly. Indonesia has articulated a national AI strategy (Badan Pengkajian dan Penerapan Teknologi [BPPT], 2020), issued an ethics circular on AI, and introduced coding and AI as an optional subject in madrasahs, yet guidance specific to Arabic pedagogy remains underdeveloped in the documents reviewed for this study.

The urgency of the question is heightened by the pace of adoption. Educational systems rarely have time to deliberate at the speed at which consumer AI products spread, and decisions taken by default, such as leaving students to use general-purpose chatbots unsupervised, are themselves pedagogical decisions with consequences. For madrasahs, whose mandate joins academic instruction with moral and religious formation, default adoption is especially problematic, because the values embedded in commercial systems were not designed with such a mandate in mind. A deliberate, principled reconceptualization is therefore preferable to a reactive one.

This article addresses these limitations through a conceptual synthesis. To reconceptualize does not mean merely to add AI tools to existing lessons, which would leave untouched the underlying assumptions about why, what, and how Arabic is taught (Kern, 2006; Selwyn, 2016). It means re-examining those assumptions in light of what AI can and cannot do. Three research questions guide the analysis: (1) What structural and pedagogical conditions characterize Arabic learning in madrasahs that AI integration must address? (2) What affordances and limitations do current AI technologies present for learning Arabic? (3) How should the purposes, content, pedagogy, teacher roles, assessment, and ethical governance of Arabic learning in madrasahs be reconceptualized?

The article makes three contributions. It consolidates dispersed literature on AI, Arabic language pedagogy, and Islamic education into a coherent analytical account; it identifies the affordances and risks of AI that are specific to Arabic; and it proposes the HUDA framework, a four-part conceptual model intended to guide policy and practice. The remainder of the article describes the library research method, presents and

discusses the findings in relation to the three research questions, and concludes with implications and directions for further research.

### Research Method

This study employs a qualitative library research design (Zed, 2004) using written sources as the primary data and adopts a narrative-critical literature review approach (Grant & Booth, 2009; Snyder, 2019) to synthesize heterogeneous scholarship and develop a conceptual framework for AI-supported Arabic language learning in madrasahs. Unlike systematic reviews that primarily estimate effects across comparable studies, this approach is intended for theory building by interpreting and integrating theoretical, empirical, and policy literature rather than evaluating the effectiveness of particular AI tools. Data were drawn from three main source categories, namely peer-reviewed journal articles, conference proceedings, and academic books on artificial intelligence in education, computer-assisted language learning (CALL), second language acquisition, Arabic language pedagogy, and Arabic natural language processing; scholarly works on Islamic educational thought, particularly *tarbiyah*, *ta'lim*, and *ta'dib*, together with the Qur'an as a normative source; and official policy documents, including Indonesian ministerial regulations, national AI strategies and ethics frameworks, and UNESCO guidance. Sources were included when they were peer-reviewed, published by recognized academic publishers, or issued by official institutions; addressed AI or educational technology in relation to language learning, Arabic language education, Islamic education, or Indonesian educational policy; and were available in English, Indonesian, or Arabic. Purely technical studies without pedagogical implications, promotional or commercial materials, and sources lacking methodological or bibliographic transparency were excluded. Policy-related claims were verified against official documents whenever possible and treated cautiously when direct verification was unavailable. Data analysis proceeded iteratively through qualitative content analysis (Krippendorff, 2018), thematic analysis following Braun and Clarke's (2006) six-phase procedure, and comparative thematic matrices informed by Miles et al. (2014) to identify convergences, tensions, and research gaps across sources. Deductive categories included learning conditions, AI affordances, risks, pedagogical models, teacher roles, and policy, while inductive codes captured concepts emerging from the literature.

## Results and Discussion

The analysis is organized around the three research questions. It first describes the conditions of Arabic learning in madrasahs (RQ1), then examines the affordances and limits of AI for Arabic (RQ2), and then presents the reconceptualization of purpose, content, pedagogy, teacher role, assessment, and governance (RQ3), culminating in an integrative framework and an implementation pathway.

### *Conditions of Arabic Learning in Madrasahs: The Baseline for Change*

The reviewed literature converges on five conditions that shape Arabic learning in madrasahs and that any integration of AI must address.

The first is a persistent tension of purpose. Arabic is taught both as a key to religious texts and as a living language of communication, and these aims pull curricula in different directions. Indonesian Islamic education has a long tradition centered on reading classical texts (*kitab kuning*) through the study of *nahwu* and *sharaf* (van Bruinessen, 1995), a tradition carried by networks of scholars across the Malay-Indonesian world and the Middle East (Azra, 2004). Modern madrasah curricula, by contrast, emphasize the four language skills in line with communicative approaches (Kementerian Agama Republik Indonesia, 2019, 2022; Richards & Rodgers, 2014). Studies of curriculum implementation suggest that teachers often negotiate between these orientations without a shared conceptual model (Ardiansyah & Muhammad, 2020).

The second condition is limited exposure. Arabic is a foreign language in Indonesian society, so learners have few opportunities to encounter it beyond school. Input-oriented and interaction-oriented theories of acquisition hold that learners need abundant comprehensible input, chances to negotiate meaning, and occasions for pushed output (Krashen, 1982; Long, 1996; Swain, 1985). Where lessons are teacher-centered and time-bound, each individual learner may speak or write Arabic for only a few minutes per lesson.

The third is diglossia. Standard Arabic (*fusha*) is the language of writing, religion, and formal media, whereas everyday communication in the Arab world takes place in regional varieties (Ferguson, 1959). Madrasah instruction teaches almost exclusively the standard variety. Some Arabic educators criticize this practice for leaving learners unprepared for authentic interaction (Al-Batal, 2018; Palmer, 2007), while others defend it on the grounds of the religious and literary purposes of madrasah study. Younes (2015)

proposes an integrated approach that treats standard and colloquial Arabic as complementary, but whether this is feasible in madrasahs remains an open question.

The fourth is heterogeneity of resources and teacher capacity. Islamic schools in Indonesia differ widely in funding, location, teacher qualifications, and affiliation with pesantren or community foundations (Hefner, 2009), and the implementation of new curricula tends to mirror this variation.

The fifth concerns assessment. Where assessment rewards the recall of rules and vocabulary, it steers teaching toward memorization irrespective of what the curriculum states, so assessment habits can quietly preserve the very practices that reform seeks to change.

These five conditions define the baseline problem to which AI is proposed as a partial solution. AI cannot resolve the tension of purpose, which requires curricular and philosophical choice, but it can address limited exposure, offer differentiated practice, and reduce some teacher workload. Conversely, AI can deepen inequity and encourage shallow processing if adopted uncritically. This dual character is developed in the next section.

### *Affordances and Limits of AI Technologies for Arabic Learning*

Four families of AI technology are most relevant to Arabic learning, summarized in Table 1. Generative chatbots and large language models offer conversational partners, on-demand explanations, text generation at controlled levels, error correction, and translation. For Arabic learners with little access to interlocutors, a chatbot can lower the affective barrier to speaking or writing and multiply practice opportunities. Arabic-specific natural language processing tools, such as morphological analyzers, diacritization systems, parsers, and pretrained models such as AraBERT, can support the analysis of word structure and syntax (Antoun et al., 2020; Habash, 2010). Speech technologies, namely automatic speech recognition and text-to-speech, can supply pronunciation models and listening materials. Adaptive learning systems and intelligent tutoring systems use learner data to sequence tasks and give step-level feedback. Meta-analytic and comparative evidence from other domains indicates that such systems can raise achievement and, under some conditions, approach the effectiveness of human tutoring (Kulik & Fletcher, 2016; VanLehn, 2011), although evidence specific to Arabic remains thin.

**Table 1** Artificial intelligence technologies relevant to Arabic learning: Affordances, limitations, and implications

| AI technology                                                                         | Pedagogical affordance for Arabic                                                                       | Principal limitations                                                                                                      | Implication for madrasah practice                                                             |
|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|
| Generative chatbots and large language models                                         | Conversation practice; on-demand explanations; graded text generation; feedback on writing; translation | Hallucination; uneven coverage of dialects and registers; overly accommodating responses; risk of replacing learner effort | Use after the learner's own attempt; teacher-designed prompts; explicit verification routines |
| Arabic NLP tools (morphological analysis, diacritization, parsing, pretrained models) | Visualization of root-pattern morphology; support for reading unvocalized texts; automated analysis     | Errors in diacritization and parsing, especially in classical or poetic texts; limited transparency                        | Treat output as a hypothesis to be checked against grammar references and the teacher         |
| Speech technologies (ASR and TTS)                                                     | Pronunciation feedback; listening materials; accessibility                                              | Reduced accuracy for learner accents and dialects; limited feedback on articulation quality                                | Use as supplementary practice; keep teacher modeling and oral interaction central             |
| Adaptive systems, intelligent tutoring, and learning analytics                        | Personalized sequencing; step-level feedback; diagnostic data on learner progress                       | Data privacy; scarce Arabic-specific evidence; cost and connectivity                                                       | Use for diagnostic and practice functions under clear data governance                         |

*Note.* Synthesized by the author from the literature reviewed in this study.

These affordances come with limitations that are especially salient for Arabic. The first is accuracy. Language models can produce confident but incorrect explanations, invent sources or quotations, and mis-vocalize words. Because Arabic script normally omits short vowels, a tool must infer them from context, and errors in diacritization can alter meaning or grammatical case (Farghaly & Shaalan, 2009; Habash, 2010). The second is uneven coverage of varieties and genres. Available Arabic text is distributed unevenly across regions, registers, and dialects, and dialectal and low-resource settings remain challenging for language technologies (Darwish et al., 2021). The third is bias and cultural fit: models reproduce the values and blind spots of their training data (Bender et al., 2021; Weidinger et al., 2021), which may not match the expectations of Indonesian Muslim learners and their teachers. The fourth is learner specificity. General-purpose systems are not designed for Indonesian speakers of Arabic, whose typical difficulties may include emphatic and pharyngeal consonants, grammatical case and gender agreement, and root-and-pattern morphology, and these may differ from the difficulties of learners for whom most tools have been developed. This last point is an inference

*Pedagogik Journal of Islamic Elementary School*

from contrastive reasoning rather than an established empirical finding and requires study.

Finally, the empirical base is uneven. Reviews of technology-enhanced language learning have long noted that studies are often short, small in scale, and methodologically heterogeneous (Golonka et al., 2014), and syntheses of AI in language education repeat this pattern (Yang & Kyun, 2022). Studies of AI in Arabic teaching are promising but exploratory. Educators should therefore treat AI as a promising but unproven set of tools and avoid sweeping claims of either transformation or harm.

### *Reconceptualizing Purpose: From Instrumental Skill to Human-Centered Literacy*

If AI can translate, vocalize, and paraphrase on demand, what is the purpose of learning Arabic in madrasahs? The analysis suggests that the answer lies in moving from an instrumental view, in which Arabic is a set of decodable skills, to a human-centered view in which Arabic is a medium of understanding, relationship, and moral formation.

Islamic educational thought supplies resources for this shift. The first revealed word of the Qur'an is the imperative اقرأ (*iqra'*, "read"), and the Qur'an speaks of the human being as taught the names of things (Abdel Haleem, 2004, 96:1; 2:31). Language is here bound to knowledge and responsibility. Al-Attas (1980) argues that the proper term for education in Islam is *ta'dib*, the inculcation of أدب (*adab*), understood as the recognition and proper placement of things within the order of knowledge and being, while Halstead (2004) shows that *tarbiyah*, *ta'lim*, and *ta'dib* together frame education as holistic development rather than the transfer of information. From this perspective, learning Arabic aims not only at competence but at a right relationship with the text, the teacher, and the truth. Delegating the labor of understanding wholesale to a machine would run against that aim.

In practical terms, the purpose can be expressed as *critical Arabic-AI literacy*, comprising four intertwined competences: (a) linguistic competence in the four skills; (b) textual and religious literacy, including the ability to read primary sources with understanding and with respect for their chains of transmission; (c) AI literacy, meaning the ability to understand what AI systems do, use them effectively, and evaluate their outputs (Holmes et al., 2019; Miao et al., 2021); and (d) ethical dispositions of honesty, humility, and responsibility in the use of knowledge. Self-determination theory adds a motivational rationale, since learners flourish when they experience autonomy, competence, and relatedness (Ryan & Deci, 2000). AI can support autonomy through

choice and personalization, but if it performs tasks on learners' behalf it may erode the sense of competence, and if it displaces teachers and peers it may weaken relatedness. The purpose of Arabic learning must therefore be stated so that AI serves the learner's own growth and does not replace it.

### *Reconceptualizing Content: Layered, Usage-Oriented Arabic*

Reconceptualizing content concerns what is taught and what learners encounter as input. Three shifts emerge from the analysis.

The first is a shift from fragmented items to language in use. Nation's (2007) four strands, namely meaning-focused input, meaning-focused output, language-focused learning, and fluency development, provide a balanced planning frame in which grammar and vocabulary study is one strand among four. Madrasah content that is heavy on language-focused learning could be rebalanced by using AI to supply additional meaning-focused input, such as graded reading texts on topics linked to learners' lives and religious practice, and additional output opportunities such as dialogue and short writing, while keeping essential grammar instruction. Vocabulary instruction can likewise benefit from frequency-based and personalized word lists (Nation, 2013). Arabic's root-and-pattern morphology, a defining feature of the language (Ryding, 2013), lends itself to AI-assisted visualizations that show how words in a family relate to one another.

The second is a shift from a single variety to layered varieties. Because conversational agents can be prompted to respond in standard Arabic or in a specified colloquial variety, they make it technically easier to expose learners to varieties beyond the textbook. This opens a curricular decision that technology cannot make: how far should madrasah programs incorporate colloquial Arabic? Arguments for integrated approaches (Al-Batal, 2018; Younes, 2015) emphasize authentic communication, whereas madrasah purposes emphasize the standard language of scripture and scholarship. A defensible position is layered. Standard Arabic (الفصحى) remains the core, with limited, teacher-selected exposure to a colloquial variety (العامية) for learners who intend to study or work in Arabic-speaking contexts. Given the uneven quality of AI-generated colloquial Arabic noted above, such exposure should be checked by teachers or proficient speakers.

The third is a shift from decoding to critical engagement with heritage texts. AI can add vowel marks, parse sentences, and paraphrase classical passages, which may

help learners approaching the *kitab kuning* tradition (van Bruinessen, 1995). Yet traditional Islamic learning does not treat a text as self-authenticating. It emphasizes reading under a teacher and transmission through a chain of authority (السند), carried by scholarly networks (Azra, 2004). This principle can be translated into a contemporary norm: AI output on classical texts is a hypothesis to be verified by the teacher and by authoritative references, never an authority in itself. Content should therefore explicitly include verification practice as a component of textual literacy.

Locally relevant content domains deserve explicit attention. For madrasah learners, meaningful use of Arabic might include the vocabulary and discourse of worship and Islamic ethics, short authentic texts drawn from Islamic sources at suitable levels, everyday and digital communication, and, for older learners, the language needed for further study or for services connected with pilgrimage and travel to Arabic-speaking countries. These domains are suggestions of the author rather than findings of a specific study, but they illustrate how a usage-oriented syllabus can be anchored in learners' actual purposes. AI can help teachers produce practice materials across such domains quickly, provided that religious content is checked against authoritative sources before it reaches learners.

### ***Reconceptualizing Pedagogy: Dialogic Human-AI Learning***

Pedagogy is where reconceptualization has the most practical consequence. Three theoretical lenses help to structure the discussion.

Sociocultural theory conceives learning as mediated by tools and social interaction, with learners progressing through the zone of proximal development when supported by more capable others (Lantolf & Thorne, 2006; Vygotsky, 1978). AI can be understood as a mediating tool that offers scaffolding, such as hints, models, and examples, at scale. Yet scaffolds must be withdrawn as competence grows, and a chatbot that always supplies the answer creates dependence rather than development. The teacher's task is to calibrate and gradually fade AI support.

Interaction and output theories suggest that chatbots can provide low-anxiety practice in producing Arabic and receiving feedback (Long, 1996; Swain, 1985). Two cautions apply. AI interlocutors tend to be accommodating and may interpret malformed input without prompting the negotiation of meaning that is thought to drive acquisition; this is an inference from the interaction approach that needs empirical testing. In addition, feedback is effective only when it is specific, related to goals, and

actionable (Hattie & Timperley, 2007), so learners need to ask AI for explanations and examples and not only for corrected text.

Cognitive load theory and the notion of desirable difficulties warn that learning requires effortful processing (Bjork, 1994; Sweller, 1988). AI that reduces extraneous load, for example by glossing unfamiliar words during reading, can help. AI that performs the productive work of composing sentences removes precisely the effort from which learning arises. This yields a first design principle, *attempt before assistance*: learners should make their own attempt before consulting AI.

Drawing on these lenses and on task-based language teaching (Ellis, 2003), five pedagogical patterns emerge as candidates for madrasah practice. The first is attempt-compare-reflect for writing (*insya'*): students draft a text, ask AI for feedback or a model version, compare and annotate the differences, and submit a short reflection. The second is constrained role-play for conversation (*muhadatsah*): the teacher sets the scenario, target vocabulary, and register, students converse with a chatbot, and the class debriefs together. The third is leveled reading with verification (*qira'ah*): AI generates or adapts texts to learners' level, followed by teacher-led checking of accuracy and discussion of meaning. The fourth is AI as a flawed informant: students analyze errors in AI-generated Arabic, building metalinguistic awareness of *nahwu* and *sharaf* as well as a critical stance toward machine output. The fifth is flipped adaptive practice, in which individualized practice takes place outside class while class time is devoted to interaction and teacher feedback. Multimodal design that pairs text with audio and images can further support comprehension (Mayer, 2009).

Two further considerations concern differentiation and affect. Madrasah classes often contain learners with very different prior exposure to Arabic, for example those who have attended pesantren or Qur'anic schools and those who have not. Adaptive practice and tiered AI-generated texts can allow each learner to work near the upper edge of current competence, provided that the teacher monitors whether tiers harden into fixed groupings. Affect matters as well. Learners who fear public mistakes may find a chatbot a lower-risk rehearsal space, although rehearsal with a machine cannot substitute for the relatedness gained by speaking with peers and teachers (Ryan & Deci, 2000). Where devices are scarce, the same patterns can be run as teacher-mediated activities, for example by projecting a single chatbot session for whole-class analysis.

These patterns should not crowd out human-to-human interaction. Traditions such as *talaqqi* and *musyafahah*, face-to-face learning in which learners hear and repeat directly from the teacher, remain central to acquiring pronunciation, and speech recognition tools should supplement rather than replace this modeling.

### ***Reconceptualizing the Teacher's Role and Professional Capacity***

The concern that AI will replace teachers finds little support in the reviewed scholarship, which consistently argues for teacher-centered integration (Holmes et al., 2019; Luckin et al., 2016; Selwyn, 2019). What changes is the nature of the teacher's work. In an AI-rich environment, Arabic teachers act as (a) learning designers who plan tasks and set boundaries for AI use; (b) curators and verifiers who select, check, and correct AI-generated material and thereby exercise linguistic authority; (c) feedback orchestrators who interpret analytics and combine AI and human feedback; and (d) moral mentors (*murabbi*) who model honesty, humility, and responsible use. The last function is irreducibly human, since the exemplary conduct (*uswah hasanah*) of the teacher is part of the pedagogy of Islamic education (Al-Attas, 1980; Halstead, 2004).

Supporting this role requires an extended view of teacher knowledge. The TPACK framework holds that effective technology integration depends on the interplay of technological, pedagogical, and content knowledge (Mishra & Koehler, 2006). For Arabic teachers in madrasahs, content knowledge includes linguistic and Islamic textual knowledge, which is exactly the knowledge needed to detect AI errors, while technological knowledge includes prompt design and awareness of model limitations. Teacher beliefs, confidence, and institutional culture shape whether such knowledge is actually put to use (Ertmer & Ottenbreit-Leftwich, 2010), and teacher education programs are already being urged to address generative AI. Professional learning is therefore best organized through sustained communities of practice (Lave & Wenger, 1991), for example within subject teacher working groups, rather than through one-off workshops.

### ***Reconceptualizing Assessment and Academic Integrity***

AI complicates assessment because take-home writing and translation can be generated in seconds, and research on second language writing notes both potential benefits and risks such as inaccuracy, plagiarism, and over-reliance. Banning AI outright is difficult to enforce and denies students legitimate learning support, while ignoring it invalidates assessments. The literature supports a redesign strategy in three parts.

First, assessment should shift toward process-oriented and performance-based formats such as drafts, oral presentations, in-class writing, portfolios, and short oral defenses of written work. This is consistent with the emphasis on diagnostic, formative, and summative assessment in the *Kurikulum Merdeka* guidelines for madrasahs (Kementerian Agama Republik Indonesia, 2022). Second, madrasahs should articulate graduated AI-use policies for each task: for instance, AI prohibited for tasks that assess unaided competence, permitted with disclosure for drafting tasks, and required for tasks that assess AI-critical skills. Third, the new competences themselves should be assessed. Rubrics should reward verification, the accurate correction of AI errors, and reflective commentary on how AI was used. Academic integrity is better protected through learning-oriented conversation than through surveillance, and teachers should be cautious about treating automated detection of AI-generated text as conclusive evidence.

Feedback literacy is a further aim. Students who learn to interpret and act on feedback, whether it comes from AI, peers, or teachers, are better placed to regulate their own learning (Hattie & Timperley, 2007). Assessment tasks can therefore ask students to compare AI feedback with teacher feedback on the same text, decide which suggestions to accept, and justify their choices by reference to grammatical rules.

### *Ethical, Epistemic, and Equity Considerations*

Five issues stand out as conditions for responsible integration.

The first is epistemic reliability, which in the Islamic context is inseparable from *tabayyun*. The Qur'an instructs believers to verify information brought by unreliable sources before acting on it (Abdel Haleem, 2004, 49:6), a principle summarized in the Indonesian scholarly vocabulary by the term *tabayyun*. This provides an indigenous ethical basis for a verification routine in classrooms. AI outputs concerning Qur'anic verses, hadith, jurisprudential rulings, and classical grammar must be checked against authoritative sources, and AI must not be treated as a source of religious rulings. Such verification echoes the scholarly value attached to *sanad* discussed earlier.

The second is bias and cultural fit. Because models learn from data that reflect particular languages, regions, and worldviews (Bender et al., 2021; Weidinger et al., 2021), their explanations and examples may embed assumptions that do not align with the values or contexts of madrasah communities. Teachers should review the examples and scenarios that AI generates for cultural and religious appropriateness.

The third is privacy and child protection. Madrasah students are mostly minors, and cross-national ethical frameworks converge on principles such as beneficence, non-maleficence, autonomy, justice, and explicability (Floridi & Cowls, 2019; Jobin et al., 2019). UNESCO's guidance on generative AI calls for data protection regulation and proposes an age threshold for independent classroom use. In Indonesia, the Personal Data Protection Law (Republik Indonesia, 2022) and the ministerial circular on AI ethics provide reference points, although the latter is addressed primarily to businesses and electronic system operators and not to schools. Madrasahs should therefore adopt institutional policies specifying which tools are approved, what data may be entered, and how parents are informed. Recent empirical work on ethical concerns in AI-supported foreign language learning reinforces the importance of such policies.

The fourth is dependence and the erosion of productive effort. If learners routinely delegate composition, translation, and grammatical analysis to AI, they may bypass the effortful processing on which durable learning depends (Bjork, 1994; Sweller, 1988). Over-reliance is also a motivational risk, since a sense of competence is one of the conditions of sustained engagement (Ryan & Deci, 2000). Guardrails include the attempt-before-assistance principle, tasks that require in-class production without AI, and periodic reflection in which learners judge which tasks they can now perform independently.

The fifth is equity. Digital inequality is a long-standing theme in the critical study of educational technology (Selwyn, 2016), and AI may widen gaps between well-resourced and under-resourced madrasahs. Mitigation strategies include teacher-mediated use on shared devices, offline-capable resources prepared in advance, priority for low-cost tools, and support from government and community partners. An equity lens also cautions against policies that presuppose every learner has a personal device and stable connectivity at home.

#### ***An Integrative Proposal: The HUDA Framework***

Integrating the findings, this article proposes the HUDA framework, an acronym for Human-centered, Usage-oriented, Dialogic, and Accountable.<sup>1</sup> Each element corresponds to a reconceptualized domain, and the teacher stands at the center as designer and exemplar. *Human-centered* refers to purpose: Arabic learning is directed at critical Arabic-AI literacy and moral formation, so AI is judged by whether it serves the

learner's growth. *Usage-oriented* refers to content: learners meet Arabic in meaningful use across the four skills, with the standard language as core, limited layered exposure to varieties, and explicit verification practice for heritage texts. *Dialogic* refers to pedagogy: learning unfolds through interaction among learners, teacher, texts, and AI, with attempt-before-assistance and gradual fading of scaffolds. *Accountable* refers to governance and assessment: AI use is transparent, verified, privacy-protective, equitable, and assessed through process-oriented tasks.

Table 2 contrasts conventional orientations with the HUDA orientation and gives indicative practices. The framework does not prescribe particular tools. It offers principles that madrasahs can test, adapt, and refine in light of their own resources and pedagogical traditions.

**Table 2** *The HUDA framework: From conventional to reconceptualized Arabic learning*

| Domain (HUDA element)                   | Conventional orientation                           | Reconceptualized orientation                                                                   | Indicative practice                                                                     |
|-----------------------------------------|----------------------------------------------------|------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|
| Purpose (Human-centered)                | Arabic as a decoding skill for texts and tests     | Critical Arabic–AI literacy and moral formation (adab)                                         | Learning goals that name linguistic, textual, AI-critical, and ethical outcomes         |
| Content (Usage-oriented)                | Isolated rules, word lists, and translation drills | Language in use; standard Arabic as core with layered exposure; verification of heritage texts | Four-strand planning; graded texts; teacher-verified colloquial samples                 |
| Pedagogy (Dialogic)                     | Teacher transmission; AI absent or used covertly   | Learner–teacher–AI interaction; attempt before assistance; fading scaffolds                    | Attempt–compare–reflect; constrained role-play; AI as flawed informant                  |
| Teacher role (cross-cutting)            | Sole source of knowledge                           | Designer, verifier, feedback orchestrator, and moral mentor                                    | Communities of practice; TPACK-informed training; shared repository of verified prompts |
| Governance and assessment (Accountable) | Recall-based tests; unregulated AI use             | Process-oriented assessment; graduated AI-use policy; privacy and equity safeguards            | Portfolios, oral defenses, disclosure statements; approved-tool list; offline options   |

*Note.* Developed by the author on the basis of the analysis presented in Sections 3.1 to 3.8.

To make the framework testable, each element can be linked to observable indicators. For the human-centered element, indicators include learning goals that name ethical and textual outcomes and evidence that learners reflect on their own use of AI. For the usage-oriented element, they include the proportion of lesson time devoted to meaning-focused input and output and the presence of verification tasks for heritage

texts. For the dialogic element, they include documented routines of attempt before assistance and observable teacher–learner–AI interaction rather than solitary use of tools. For the accountable element, they include a written AI-use policy, disclosure practices in assessment, approved-tool lists, and equitable access arrangements. Such indicators could serve as the basis for classroom observation instruments, self-evaluation tools for madrasahs, and future empirical research.

Translating the framework into practice suggests a phased pathway rather than a single rollout.

- **Phase 1, Foundation:** audit infrastructure, teacher competences, and current policies; adopt an institutional AI-use policy; and provide baseline training in AI literacy and verification routines.
- **Phase 2, Pilot:** implement selected patterns, such as attempt–compare–reflect, in a limited number of classes, using action research or design-based inquiry to gather evidence on learning, motivation, and integrity.
- **Phase 3, Institutionalization:** integrate validated practices into lesson plans and assessment, build a shared repository of verified prompts and materials, and network across madrasahs.

At the system level, the Ministry of Religious Affairs and university departments of Arabic education could co-develop Arabic-specific guidance, vetted resources, and research partnerships, building on the existing curriculum framework and the optional AI subject. Such coordination would help avoid a pattern in which each school independently rediscovers the same risks.

Policy alignment is also needed. Because curriculum guidelines for madrasahs set learning outcomes and organize teacher planning (Kementerian Agama Republik Indonesia, 2022), the most direct route to influence is guidance that shows teachers how AI-supported activities can serve those outcomes, for example through sample lesson plans and assessment rubrics that make the expected human and machine contributions explicit. Such guidance should be provisional and revisable, given the pace of technological change, and should be developed with practicing teachers so that it reflects the realities of classrooms with limited devices and uneven connectivity.

## Conclusion

This article has argued that AI should prompt a reconceptualization, and not merely a technological upgrade, of Arabic language learning in madrasahs. In response

*Pedagogik Journal of Islamic Elementary School*

to the first research question, the literature describes a baseline of tension between textual and communicative purposes, limited exposure to Arabic, diglossia, uneven resources, and assessment habits that may favor recall. In response to the second, AI offers valuable affordances, including conversation practice, adaptive feedback, graded input, and morphological analysis, but its limitations are acute for Arabic: hallucination, uneven coverage of varieties, script-related ambiguity, bias, and a thin evidence base. In response to the third, the analysis identified four interrelated shifts in purpose, content, pedagogy, and governance, which are integrated in the HUDA framework with the teacher at its center.

Several implications follow. For teachers, the priority is to develop the combined linguistic, pedagogical, and technological knowledge needed to design AI-supported tasks, to verify AI output, and to model ethical use. For madrasah leaders, the priority is to establish clear AI-use policies, invest in collaborative professional learning, and attend to equity in access. For policymakers, the priority is to translate general AI ethics and curriculum policies into guidance specific to Arabic instruction and to support the development of vetted resources. For researchers, the priority is empirical work, including quasi-experimental and design-based studies of the pedagogical patterns proposed here, analyses of how AI handles the errors of Indonesian learners of Arabic, studies of teacher cognition, and research on equity and access.

The proposals in this article should be read as invitations to experiment responsibly, not as mandates. Madrasahs differ, and the most defensible path in many contexts will begin small: a single class, a single pedagogical pattern, careful documentation of what happens, and open discussion among teachers about what was learned.

The study has limitations. It is a conceptual synthesis and does not provide empirical proof of the framework's effects; the rapid evolution of AI means that specific claims about tool performance will need updating; and the literature on Arabic and on madrasahs is smaller than that on English-language contexts. These limitations mark the framework as a proposition to be tested rather than a settled model.

Technology will continue to change, but the enduring purposes of learning Arabic in madrasahs, namely understanding the revealed text, sustaining a scholarly heritage, communicating across communities, and cultivating *adab*, do not change with

it. The task for Arabic educators in the age of AI is to ensure that increasingly capable machines serve these purposes and do not quietly redefine them.

## References

- Abdel Haleem, M. A. S. (Trans.). (2004). *The Qur'an*. Oxford University Press.
- Al-Attas, S. M. N. (1980). *The concept of education in Islam: A framework for an Islamic philosophy of education*. Muslim Youth Movement of Malaysia.
- Al-Batal, M. (Ed.). (2018). *Arabic as one language: Integrating dialect in the Arabic language curriculum*. Georgetown University Press.
- Antoun, W., Baly, F., & Hajj, H. (2020). AraBERT: Transformer-based model for Arabic language understanding. In *Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools* (pp. 9–15). European Language Resources Association. <https://aclanthology.org/2020.osact-1.2>
- Ardiansyah, A. A., & Muhammad, A. (2020). Implementation of integrative Arabic grammar (Nahwu & Sharaf) curriculum in Islamic boarding school. *Izdihar: Journal of Arabic Language Teaching, Linguistics, and Literature*, 3(3), 211–228. <https://doi.org/10.22219/jiz.v3i3.13264>
- Azra, A. (2004). *The origins of Islamic reformism in Southeast Asia: Networks of Malay-Indonesian and Middle Eastern 'ulama' in the seventeenth and eighteenth centuries*. Allen & Unwin; University of Hawai'i Press.
- Badan Pengkajian dan Penerapan Teknologi. (2020). *Strategi nasional kecerdasan artifisial Indonesia 2020–2045* [Indonesia's national strategy for artificial intelligence 2020–2045].
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445922>
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). MIT Press.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. In *Advances in Neural Information Processing Systems* (Vol. 33, pp. 1877–1901). Curran Associates.
- Darwish, K., Habash, N., Abbas, M., Al-Khalifa, H., Al-Natsheh, H. T., Bouamor, H., Bouzoubaa, K., Cavalli-Sforza, V., El-Beltagy, S. R., El-Hajj, W., Jarrar, M., & Mubarak, H. (2021). A panoramic survey of natural language processing in the Arab world. *Communications of the ACM*, 64(4), 72–81. <https://doi.org/10.1145/3447735>

- Effendy, A. F. (2005). *Metodologi pengajaran bahasa Arab*. Misykat.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs, and culture intersect. *Journal of Research on Technology in Education*, 42(3), 255–284. <https://doi.org/10.1080/15391523.2010.10782551>
- Farghaly, A., & Shaalan, K. (2009). Arabic natural language processing: Challenges and solutions. *ACM Transactions on Asian Language Information Processing*, 8(4), Article 14. <https://doi.org/10.1145/1644879.1644881>
- Ferguson, C. A. (1959). Diglossia. *Word*, 15(2), 325–340. <https://doi.org/10.1080/00437956.1959.11659702>
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27(1), 70–105. <https://doi.org/10.1080/09588221.2012.700315>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Habash, N. Y. (2010). *Introduction to Arabic natural language processing*. Morgan & Claypool. <https://doi.org/10.2200/S00277ED1V01Y201008HLT010>
- Halstead, J. M. (2004). An Islamic concept of education. *Comparative Education*, 40(4), 517–529. <https://doi.org/10.1080/0305006042000284510>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hefner, R. W. (Ed.). (2009). *Making modern Muslims: The politics of Islamic education in Southeast Asia*. University of Hawai'i Press.
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Kementerian Agama Republik Indonesia. (2019). *Keputusan Menteri Agama Republik Indonesia Nomor 183 Tahun 2019 tentang Kurikulum Pendidikan Agama Islam dan Bahasa Arab pada Madrasah* [Decree of the Minister of Religious Affairs No. 183 of 2019 on the Islamic religious education and Arabic curriculum in madrasahs].
- Kementerian Agama Republik Indonesia. (2022). *Keputusan Menteri Agama Republik Indonesia Nomor 347 Tahun 2022 tentang Pedoman Implementasi Kurikulum Merdeka pada Madrasah* [Decree of the Minister of Religious Affairs No. 347 of 2022 on guidelines for implementing the Independent Curriculum in madrasahs].
- Kern, R. (2006). Perspectives on technology in learning and teaching languages. *TESOL Quarterly*, 40(1), 183–210. <https://doi.org/10.2307/40264516>
- Pedagogik Journal of Islamic Elementary School*

- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications.
- Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, 86(1), 42–78. <https://doi.org/10.3102/0034654315581420>
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.
- Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: Guidance for policy-makers*. UNESCO.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Nation, I. S. P. (2007). The four strands. *Innovation in Language Learning and Teaching*, 1(1), 2–13. <https://doi.org/10.2167/illt039.0>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Palmer, J. (2007). Arabic diglossia: Teaching only the standard variety is a disservice. *Arizona Working Papers in SLA & Teaching*, 14, 111–122.
- Republik Indonesia. (2022). *Undang-Undang Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi* [Law No. 27 of 2022 on personal data protection].
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge University Press.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Ryding, K. C. (2013). *Teaching and learning Arabic as a foreign language: A guide for teachers*. Georgetown University Press.
- Selwyn, N. (2016). *Education and technology: Key issues and debates* (2nd ed.). Bloomsbury Academic.
- Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.

- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. [https://doi.org/10.1207/s15516709cog1202\\_4](https://doi.org/10.1207/s15516709cog1202_4)
- van Bruinessen, M. (1995). *Kitab kuning, pesantren dan tarekat: Tradisi-tradisi Islam di Indonesia*. Mizan.
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221. <https://doi.org/10.1080/00461520.2011.611369>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 5998–6008). Curran Associates.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.
- Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., Kenton, Z., Brown, S., Hawkins, W., Stepleton, T., Biles, C., Birhane, A., Haas, J., Rimell, L., Hendricks, L. A., ... Gabriel, I. (2021). *Ethical and social risks of harm from language models* (arXiv:2112.04359). arXiv. <https://doi.org/10.48550/arXiv.2112.04359>
- Yang, H., & Kyun, S. (2022). The current research trend of artificial intelligence in language learning: A systematic empirical literature review from an activity theory perspective. *Australasian Journal of Educational Technology*, 38(5), 180–210. <https://doi.org/10.14742/ajet.7492>
- Younes, M. (2015). *The integrated approach to Arabic instruction*. Routledge.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16, Article 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zed, M. (2004). *Metode penelitian kepustakaan*. Yayasan Obor Indonesia.