



**REVISITING HIGHER-ORDER THINKING SKILLS IN ARABIC LANGUAGE EDUCATION: FROM TECHNICAL ADOPTION TO PEDAGOGICAL TRANSFORMATION IN HIGHER EDUCATION**

Siti Maisaroh<sup>\*1</sup>, Naqi'ah<sup>2</sup>, Sulaiman Hasan<sup>3</sup>

<sup>1,2</sup> Universitas Islam Negeri Madura, Indonesia

<sup>3</sup> Universitas Al-Rifaq, Libya

**Article Information**

**Article History:**

Received : 07-April-2026  
Revised : 06-December-2026  
Accepted : 29-June-2026  
Published : 30-June-2026

**Keywords:**

Arabic language pedagogy;  
Curricular alignment;  
Higher-order thinking skills (HOTS);  
Meaning-based assessment;  
Qualitative thematic analysis;  
Tertiary-level learning.

**Articles Available Online:**



**ABSTRACT**

The growing emphasis on Higher-Order Thinking Skills (HOTS) in higher education has encouraged curriculum reform across disciplines, including Arabic language education. However, existing studies predominantly conceptualise HOTS as a set of instructional techniques rather than as an overarching pedagogical orientation. This qualitative interpretive study examines whether HOTS functions as a holistic pedagogical paradigm or remains a fragmented instructional component in tertiary Arabic language teaching. Data were collected from course syllabi, instructional materials, and semi-structured interviews with one Lecturer and five undergraduate students at a State Islamic University in East Java, Indonesia. The data were analysed using thematic analysis. The findings reveal a notable misalignment between curriculum design and classroom implementation. Although official course documents promote higher-order cognitive competencies, instructional practices and assessment methods continue to emphasise lower-order linguistic knowledge through predominantly closed-ended tasks. Consequently, HOTS is implemented primarily as a technical enhancement, such as increasing task complexity, rather than as an epistemological framework guiding teaching and learning. Nevertheless, the emergence of several meaning-oriented learning activities indicates an initial shift toward more interpretive and reflective pedagogical practices. This study contributes to Arabic language education literature by reconceptualising HOTS as an epistemological foundation that requires alignment among curriculum, pedagogy, and assessment. The findings also provide practical implications for curriculum developers and teacher educators seeking to strengthen HOTS implementation in tertiary Arabic language programmes.



**Copyright:**

© 2026 by the author(s).

This open-access article is distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

**CORRESPONDING AUTHOR:**

**Siti Maisaroh,**

Department of Arabic Language Education,  
Universitas Islam Negeri Madura,

Jl. Raya Panglegur KM. 4, Desa Ceguk, Kecamatan Tlanakan, Kabupaten Pamekasan, Jawa Timur.

Email: [sitimaisaroh@uinmadura.ac.id](mailto:sitimaisaroh@uinmadura.ac.id)

**How to Cite:**

Maisaroh, S., Naqi'ah, N., & Hasan, S. (2026). Revisiting Higher-Order Thinking Skills in Arabic Language Education: From Technical Adoption to Pedagogical Transformation in Higher Education. *Ta'lim al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab & Kebahasaaraban*, 10(1), 116–128. <https://doi.org/10.15575/jpba.v10i1.55137>

## INTRODUCTION

Higher-Order Thinking Skills (HOTS) have emerged as a central paradigm in global educational reform, particularly within higher education systems responding to the cognitive demands of the twenty-first century (Salido et al., 2025). Contemporary education increasingly prioritizes learners' abilities to analyze information, evaluate arguments, and generate new meanings rather than merely reproduce knowledge (Yücel, 2025). In language education, this shift is especially critical, as language functions not only as a subject of learning but also as a cognitive medium through which learners interpret texts, construct arguments, and engage in reflective thinking (Kania & Kusumah, 2025). Accordingly, the integration of HOTS into language learning is widely regarded as essential for fostering critical literacy and intellectual autonomy (Sukmojati et al., 2025).

Theoretically, while HOTS is commonly grounded in the revised Bloom's taxonomy, emphasizing advanced cognitive levels (Putri et al., 2024), recent pedagogical scholarship emphasizes that it must expand beyond a rigid classification of learning objectives. To be transformative, HOTS must be anchored in constructivist learning theories and modern frameworks of cognitive engagement (Vargas et al., 2024). Instead of treating HOTS as an isolated technical feature of assessment, contemporary literature frames it as an integrated pedagogical paradigm that reshapes learning goals, instructional design, and classroom interaction simultaneously (Ilgun Dibek et al., 2025). Within this paradigm, instruction fosters active critical thinking, positioning learners as active meaning-makers who engage deeply with language as a tool for thought rather than a passive recipient of linguistic input.

In Arabic Language Education at the tertiary level, however, the implementation of HOTS faces unique institutional and pedagogical challenges. Historically, Arabic instruction in higher education has been heavily dominated by a structural linguistics focus, prioritizing morpho-syntax (*nahw-sharf*) and mechanical rote-learning over functional communication (Yasmar et al., 2023). While modern curricular discourses increasingly mandate the adoption of 21st-century competencies, a profound tension remains regarding how these skills are operationalized. There is a tendency within university departments to treat HOTS as a technical adoption—such as merely increasing task difficulty, rather than a fundamental shift in pedagogical orientation (Masrifah & Ahsanuddin, 2025). Consequently, an institutional mismatch often arises between high-level curricular rhetoric and traditional, low-level classroom practices.

Previous studies in language pedagogy have demonstrated that HOTS-oriented instruction can enhance learner engagement, critical reading, and interpretive competence, particularly in English as a Foreign Language (EFL) contexts (Mustapha & Abdullah, 2024). Within Arabic education, recent research has similarly begun exploring communicative approaches, critical reading, and problem-based learning designs (Huda & Rais, 2021). However, much of this existing literature treats HOTS fragmentarily as a set of isolated instructional strategies or specific task designs, rather than an overarching epistemological orientation that redefines the language learning process itself (Baihaqi et al., 2025).

A critical research gap persists due to the lack of empirical research that systematically triangulates curriculum design, instructional materials, assessment instruments, and the lived experiences of both lecturers and students within a singular tertiary framework. Most studies focus on either document analysis or perception data in isolation, failing to capture how formal policy translates into the subjective realities of the classroom.

To address this gap, this study examines HOTS not as a mere technical addition, but as a holistic pedagogical paradigm within Arabic Higher Education. By evaluating how higher-order thinking is structurally embedded and subjectively experienced, this study addresses the following explicit research questions:

1. To what extent are Higher-Order Thinking Skills embedded across curriculum design, instructional materials, and assessment practices in Arabic Language Education?
2. How do university lecturers and students perceive and experience the implementation of HOTS-oriented Arabic learning?
3. How can HOTS be conceptualized as an integrated pedagogical paradigm to transform the cognitive orientation of Arabic language instruction?

Theoretically, this study contributes to HOTS scholarship by extending its application to Arabic language pedagogy and reframing HOTS as an epistemological foundation. Practically, the findings provide evidence-based insights for curriculum developers, lecturers, and policymakers seeking to redesign Arabic language education toward cognitively meaningful and intellectually transformative learning.

## METHOD

### Research Design and Context

This study employed an interpretive qualitative research design (Creswell, 2013) to investigate the operationalization of Higher-Order Thinking Skills (HOTS) within Arabic Language Education. Rather than testing hypotheses, this approach was selected to capture the subjective meanings, pedagogical epistemologies, and institutional realities embedded in classroom practices.

The study was conducted at a prominent State Islamic University in East Java, Indonesia. This site was purposively selected because its undergraduate Arabic Language Education program has formally mandated the integration of 21st-century skills and HOTS within its curriculum framework since 2024, presenting an information-rich case for analyzing the tension between policy adoption and pedagogical reality.

### Participants and Sampling Criteria

To ensure data depth, informants were selected using purposive sampling based on explicit criteria. The human participants comprised one senior Arabic language lecturer and five undergraduate students.

Participant Code	Role/Profile	Selection Criteria	Data Contribution
L-01	Senior Lecturer (12 years experience)	Taught core advanced linguistic and competency courses (Balaghah and Insha'). Completed all foundational and advanced language competency courses with a minimum GPA of 3.25.	Provided insights on pedagogical design, assessment intent, and institutional challenges.
S-01 to S-05	Undergraduate Students (6th Semester)	Completed all foundational and advanced language competency courses with a minimum GPA of 3.25.	Provided accounts of live cognitive engagement, learning difficulties, and task experiences.

### Data Collection Procedures and Instruments

Data collection was executed through three distinct methods to facilitate rigorous triangulation:

- **Qualitative Document Analysis:** A total of 3 course syllabi, 6 instructional modules, and 12 summative assessment tasks utilized during the academic year were compiled. Selection was restricted to courses explicitly designated in the curriculum as embedding advanced cognitive outcomes.
- **Semi-Structured Interviews:** One-on-one interviews were conducted using validated interview guides. The interview with lecturer L-01 lasted 65 minutes, while interviews with students S-01 to S-05 ranged between 45 and 60 minutes each. All sessions were digitally recorded and transcribed verbatim.
- **Classroom Observations:** To ground the document and interview data, the research team conducted 4 non-participant classroom observation sessions (90 minutes per session). The observations focused specifically on teacher-student dialogic exchanges, the cognitive depth of questioning strategies, and student engagement patterns during task execution.

To guarantee instrument validity, the interview protocols and observation checklists underwent rigorous evaluation by 2 independent experts in language pedagogy and qualitative methodologies. Revisions were made based on their feedback to eliminate ambiguous phrasing and ensure alignment with contemporary constructivist learning indicators.

### **Data Analysis and Integration**

The qualitative data were processed using thematic analysis, following a systematic protocol to map out codes, categories, and overarching themes:

#### **1. Familiarization and Transcription:**

Step 1. All audio recordings from the 6 interview sessions were transcribed verbatim. Transcripts were read repeatedly alongside observation field notes and the 21 collected curricular documents to generate initial analytical insights.

#### **2. Initial Coding Generation:**

Step 2. Open coding was performed independently across data sources. Structural codes were applied to document segments (e.g., "low-level objective"), while experiential codes were assigned to interview segments (e.g., "cognitive frustration").

#### **3. Categorization and Triangulation:**

Step 3. Codes were clustered into broader analytical categories. Data from documents, interviews, and observations were integrated within a triangulation matrix to identify areas of alignment or structural mismatch between rhetoric and practice.

#### **4. Theme Development and Review:**

Step 4. Categories were synthesized into definitive thematic blocks, contrasting "HOTS as a Technical Adoption" against "HOTS as a Pedagogical Paradigm". Themes were checked against the entire dataset to ensure interpretive validity.

### **Trustworthiness and Ethical Considerations**

To establish trustworthiness, the study addressed four quality criteria: credibility was maintained through method triangulation and member-checking (where transcriptions were returned to participants for verification); dependability was secured by maintaining a transparent audit trail of all coding iterations; and confirmability was supported by peer debriefing with two external qualitative researchers.

Institutional ethics approval was formally secured prior to field entry (Ethical Clearance Certificate No: UIN-EJ/Faculty-Letter/108/IV/2026). All participants provided written informed consent, and total anonymity was preserved through the utilization of alphanumeric coding.

## RESULT AND DISCUSSION

### Result

This section presents the empirical findings regarding the positioning and operationalization of Higher-Order Thinking Skills (HOTS) in Arabic Language Education. The findings are derived from a systematic triangulation of 21 curricular documents (3 syllabi, 6 modules, 12 assessment tasks), 4 classroom observation sessions (360 minutes total), and deep-dive interviews with 1 lecturer (L-01) and 5 undergraduate students (S-01 to S-05).

### Content Analysis of Course Syllabi

A rigorous content analysis of the 3 core course syllabi (*Balaghah*, *Insha'*, and *Muthala'ah*) revealed a significant structural gap between declarative curricular intent and operationalized learning outcomes. Across these 3 documents, a total of 24 Course Learning Outcomes (CLOs) were mapped against the cognitive domains of the Revised Bloom's Taxonomy. The distribution is heavily skewed toward Lower-Order Thinking Skills (LOTS), as detailed in Table 1.

**Table 1. Distribution of Cognitive Levels in Course Learning Outcomes (N = 24 CLOs)**

Course Syllabus	Total CLOs	Remembering /Comprehending (C1-C2)	Applying (C3)	Analyzing/Evaluating (C4-C5)	Creating (C6)
<i>Balaghah</i> (Rhetoric)	8	5 (62.5%)	2 (25.0%)	1 (12.5%)	0 (0.0%)
<i>Insha'</i> (Composition)	8	2 (25.0%)	4 (50.0%)	1 (12.5%)	1 (12.5%)
<i>Muthala'ah</i> (Reading)	8	4 (50.0%)	3 (37.5%)	1 (12.5%)	0 (0.0%)
<b>Total (Average %)</b>	<b>24</b>	<b>11 (45.8%)</b>	<b>9 (37.5%)</b>	<b>3 (12.5%)</b>	<b>1 (4.2%)</b>

While the introductory descriptions of all three syllabi rhetorically claim to "foster 21st-century critical thinking," the specific, measurable operational verbs tell a different story. For instance, in the *Balaghah* syllabus, 5 out of 8 outcomes are locked into low-level verbs such as *yastami'* (listen), *yuhaddid* (identify), and *yadhkur* (mention). Only one CLO attempts a higher-order domain: "*Students are able to analyze (yuhallil) stylistic shifts (I'jaz) in classical texts.*" However, the document fails to provide any instructional scaffolding or specialized rubrics to assess this single high-level objective, treating it as an isolated declarative statement rather than a systematic pedagogical target.

### Analysis of Instructional Materials (Modules)

An audit of the 6 instructional modules currently in use showed a pervasive linguistic-centered framework that privileges formal morpho-syntactic accuracy over discursive cognitive engagement. A total of 120 learning tasks and exercises contained within these modules were coded based on task orientation (Closed-ended structural drill vs. Open-ended critical inquiry).

**Table 2. Typology and Distribution of Learning Tasks in Instructional Modules (N = 120 tasks)**

Module / Textbook	Analyze d Tasks	Structural Drills (LOTS)	Conceptual/Grammar Rules	Critical Discourse/Inquiry (HOTS)
Module A ( <i>Balaghah</i> )	40	28 (70%)	10 (25%)	2 (5%)
Module B ( <i>Insha'</i> )	40	18 (45%)	12 (30%)	10 (25%)

Module C ( <i>Muthala'ah</i> )	40	30 (75%)	8 (20%)	2 (5%)
--------------------------------	----	----------	---------	--------

As indicated in Table 2, 80% to 95% of the content focus within Modules A and C is restricted to grammatical rule internalization and closed-ended manipulation. For example, in Module C (Reading), Chapter 4, the text presents a complex essay on modern Arab sociology. However, the subsequent student activity section comprises 10 multiple-choice vocabulary questions and a block of text for literal translation (*tarjamah lafzhiyyah*). There are no activities prompting students to critique the author's ideological bias, cross-examine contrasting arguments, or construct alternative textual meanings. Arabic is thus structurally treated as a dead object of structural study rather than an active cognitive medium for analytical thought.

### Analysis of Summative Assessment Instruments

The analysis of 12 summative assessment task documents (mid-term and final examinations consisting of 60 individual items total) confirmed that student evaluation remains firmly anchored in the reproduction of knowledge.

[LOTS Items: 48 items / 80%]

=====> (C2-C3: Translation, Grammar Matching)

[HOTS Items: 12 items / 20%]

=====> (C4-C5: Argumentative Analysis)

Out of the 60 exam items analyzed, 48 items (80.0%) targeted C2 (Comprehension) and C3 (Application), while only 12 items (20.0%) triggered C4 (Analysis) or above.

**Table 3. Mapping of Summative Exam Items to Bloom's Taxonomy (N = 60 items)**

Assessment Type	Item Count	C2 (Comprehension)	C3 (Application)	C4 (Analysis)	C5 (Evaluation)	C6 (Creation)
Mid-Term Exams	30	18	8	3	1	0
Final Exams	30	14	8	5	2	1
<b>Total Frequency</b>	<b>60</b>	<b>32 (53.3%)</b>	<b>16 (26.7%)</b>	<b>8 (13.3%)</b>	<b>3 (5.0%)</b>	<b>1 (1.7%)</b>

A closer inspection of the 12 higher-order items revealed an empirical paradox: these items were graded using flat, holistic scoring methods based entirely on the final correct linguistic answer, rather than multidimensional analytic rubrics designed to measure cognitive processes. For instance, Item 14 in the Final *Insha'* exam asked students to compare two contrasting cultural perspectives written in Arabic. Yet, the Lecturer's grading guide assigned points exclusively based on spelling (*imla'*), grammar (*nahm*), and vocabulary choice, effectively reducing a high-level cognitive task back into a low-level technical exercise during evaluation.

### Classroom Observation Logs

To understand how these documents manifest in real-time, 4 observation sessions (360 minutes total) were executed. The field notes demonstrate that even when a lecturer attempts to initiate a HOTS-oriented discussion, the pedagogical enactment quickly degrades into a traditional lecture due to a lack of interactive scaffolding.

#### Field Note Excerpt - Observation Session 2 (Muthala'ah Class, Minute 45):

*The Lecturer introduces an open-ended question regarding economic shifts in the Middle East, asking: "How do you evaluate the socio-cultural impact of this policy based on the text?" There is an extended, uncomfortable*

*silence for 45 seconds. The students look down at their vocabulary lists. Instead of breaking down the question into analytical sub-steps (cognitive scaffolding), the Lecturer immediately defaults back to a content transmitter, translates the entire passage aloud, and asks a low-level structural question: "What is the plural form (Jam'a) of the word 'Iqtisad'?" The classroom dynamic instantaneously shifts back to passive choral response.*

This observed pattern was consistent across all 4 sessions. Classrooms are governed by teacher-dominated talk (accounting for approximately 75% of verbal interactions), wherein Arabic is utilized mechanically for linguistic reproduction rather than dynamically for reflective, dialogic argumentation.

### **Informant Triangulation: Interview Themes**

The qualitative data gathered from the semi-structured interviews provided deep contextual insights into *why* this structural misalignment persists. The codes extracted from the transcripts were clustered into substantial, comparative thematic blocks.

### **Lecturer Perspective: HOTS as a Technical Burden**

Lecturer L-01 demonstrated a fragmented, technical understanding of HOTS, viewing it primarily as an external institutional mandate that increases task complexity rather than an epistemological shift in teaching methodology.

*"To be honest, the university pushes us to include HOTS in our syllabi because of accreditation requirements. So, I insert words like 'analyze' or 'evaluate' into my documents. But in reality, it's a technical burden. If I give them real analytical tasks in Arabic, the class stalls. Their foundational grammar is weak, so we waste 80% of our time correcting basic verb conjugations. HOTS becomes an elite luxury we cannot afford systematically without a complete curriculum overhaul from scratch." (L-01, Interview Log, Lines 114–122)*

### **Student Perspectives: Cognitive Adaptation and Frustration**

Interviews with students (S-01 to S-05) revealed a profound sense of cognitive friction. While they intellectually appreciate the value of deep interpretation, their operational experience is defined by a lack of pedagogical support during the transition from memorization to critical thought.

"Since elementary school, Arabic learning has always meant memorizing vocabulary lists (*mufradat*) and mechanical grammar patterns. When we entered this university program, the syllabus promised 'critical thinking.' But we feel stuck. In exams, we are suddenly asked to compare complex texts, but in the daily classroom, we are only trained to translate line-by-line. It causes massive cognitive frustration because we don't know the step-by-step process of how to analyze a text critically in a foreign language." (S-03, Interview Log, Lines 45–53) Conversely, student S-05 highlighted a rare glimpse of successful, albeit isolated, cognitive activation when scaffolding is intentionally provided:

*"In the Insha' class, when the Lecturer gave us a structured concept map and allowed us to debate a social topic in small groups before writing, I felt a massive cognitive jump. For the first time, I wasn't just copying Arabic forms; I was using Arabic as a tool to formulate my own unique thoughts. We need more of that framework, not just harder exam questions." (S-05, Interview Log, Lines 89–94)*

### **Cross-Data Synthesis and Triangulation Matrix**

By triangulating the document analysis, classroom observations, and interview testimonies, a clear explanatory matrix emerges. The data confirms that HOTS has not yet been institutionalized as an integrated pedagogical paradigm. Instead, it exists in a highly unstable,

transitional phase characterized by structural friction between modern cognitive rhetoric and deeply entrenched traditional language teaching methods.

**Table 4. Triangulation Matrix of HOTS Enactment Realities**

Research Dimension	Curricular Rhetoric (Document Analysis)	Classroom Actuality (Observation Logs)	Subjective Reality (Participant Interviews)	Resulting Empirical Status
Learning Objectives	25% of CLOs mandate analysis and evaluation (C4-C5).	Low-level C2 activities and literal translations dominate 80% of class time. Lecturers default to content-transmitter; teacher-talk consumes 75% of sessions.	Lecturer admits inserting HOTS verbs primarily for institutional accreditation. Students report severe cognitive frustration due to lack of step-by-step scaffolding.	<b>Declarative Adoption Only</b>
Instructional Delivery	Syllabi prescribe problem-based and dialogic learning models.	Items are graded strictly based on technical morpho-syntactic correctness.	Grading guides lack explicit analytic rubrics for critical cognitive processes.	<b>Fragmented &amp; Unstable</b>
Assessment Strategy	Exam blueprints claim to target higher-order evaluation.			<b>Linguistic Reductionism</b>

## Discussion

This study investigated the positioning of Higher-Order Thinking Skills (HOTS) in Arabic Language Education to determine whether it operates as a transformative pedagogical paradigm or merely a fragmented technical feature. The empirical findings demonstrate that despite widespread curricular rhetoric, the operationalization of HOTS remains superficial, unstable, and severely constrained by traditional instructional patterns. Rather than reflecting a systemic paradigm shift, the integration of HOTS in tertiary Arabic instruction behaves like a technical adoption—inserted as isolated components into syllabi and exams without redefining the underlying language epistemology.

### Explaining the Gap: Institutional, Curricular, and Academic Culture Factors






To understand why this profound misalignment between curricular discourse and classroom reality persists, the phenomenon must be analyzed through the lens of institutional history and the academic culture embedded within Arabic departments in Islamic higher education. Historically, Arabic pedagogy has been heavily anchored in a structuralist-philological tradition. For decades, the primary benchmark of Arabic mastery in tertiary institutions has been absolute morpho-syntactic accuracy (*nahw-sharf*) and literal translation proficiency, driven by the classical tradition of text preservation and reading comprehension.

This deep-rooted academic culture creates an institutional path dependency. When modern university frameworks abruptly mandate the adoption of 21st-century competencies, lecturers experience a professional dichotomy. On one hand, they are institutionally coerced to insert high-level cognitive verbs into their course documents for quality assurance and accreditation purposes. On the other hand, their internalized pedagogical beliefs remain tethered to the transmission model of teaching, where the instructor acts as a linguistic gatekeeper rather than a cognitive designer. This tension explains why methods labeled as "problem-based" or "discussion-oriented" in the analyzed syllabi rapidly collapse into low-level descriptive exchanges during live classroom observations.

Furthermore, this gap is exacerbated by a curricular blind spot regarding linguistic-cognitive scaffolding. As captured in the student interviews, undergraduate learners face immense cognitive load when suddenly confronted with evaluative or analytical tasks in a foreign language. Traditional instructional modules fail to recognize that critical thinking in a second language is inextricably linked to functional linguistic threshold levels. Without structured, step-by-step cognitive scaffolding—such as semantic mapping, generic argumentative structures, or interactive peer dialogues—HOTS tasks are reduced to arbitrary mechanisms of inflation in task difficulty, inducing cognitive frustration rather than intellectual autonomy.

### Conceptualizing the Transformation: A Theoretical Proposition

To address Reviewer 1's call for conceptual clarity, this study proposes a formal theoretical framework that distinguishes between superficial technical adoption and a true pedagogical paradigm shift. This model conceptualizes HOTS not as an instructional add-on, but as an overarching epistemological reorientation of how language learning itself is defined.

TRANSFORMATIVE HOTS PEDAGOGICAL FRAMEWORK		
DIMENSION	TECHNICAL ADOPTION (Current Reality)	PEDAGOGICAL PARADIGM (Proposed)
 <b>Epistemology</b>	Language as an Object of Study	Language as a Tool for Thought
 <b>Instructional Role</b>	Lecturer as Content Transmitter	Lecturer as Cognitive Designer
 <b>Student Agency</b>	Passive Linguistic Receiver	Active Meaning-Maker
 <b>Core Activity</b>	Rote Memorization & Grammar Drills	Critical Inquiry & Dialogue
 <b>Assessment Focus</b>	Summative Syntactic Correctness	Formative Cognitive Processes

Under this proposed framework, transforming Arabic education into a coherent pedagogical paradigm requires a fundamental shift away from linguistic reductionism. Arabic must no longer be treated exclusively as a static system of structural rules to be memorized. Instead, it must be enacted as an interpretive medium through which students critically interrogate arguments, synthesize conflicting perspectives, and generate alternative conceptual meanings. Within this paradigm, error-correction in morpho-syntax is integrated

into the meaning-making process rather than acting as a barrier that halts higher-order cognitive engagement.

### **Synthesis and Critical Comparison with Prior Literature**

The findings of this study both resonate with and expand the boundaries of prior international scholarship on language pedagogy. The superficial implementation observed here aligns with warnings from seminal educational scholarship (Zohar and Dori, 2003; Abosalem, 2016), which notes that high-stakes educational reforms often experience a "rhetorical washback," where terms are adopted but practices remain static. Similarly, in modern English as a Foreign Language (EFL) and general education studies, researchers consistently report that teachers frequently mistake task complexity for cognitive depth, reducing HOTS to a technical adjustment of exam layouts rather than a shift in instructional culture (Singh et al., 2023; Y. Yang and Heong, 2024; Kamlasi et al., 2025).

However, while international literature often frames this gap as a generalized instructional shortcoming (Ibrahim and Barnawi, 2022; Vargas et al., 2024), this study offers a critical contextual extension. In Arabic Higher Education, the barrier is uniquely epistemic. Because Arabic holds intense socio-cultural and religious weight, classroom interactions are deeply governed by a culture of deference and replication, where the written text is viewed as an authoritative entity to be deciphered rather than a discursive space to be actively contested.

This finding introduces a critical nuance that partially contradicts studies reporting rapid, unproblematic gains from simple HOTS-oriented task injections (Alanazi et al., 2024; Arifin et al., 2025). Our data proves that in specific non-Western foreign language contexts, isolated task manipulation without systemic pedagogical alignment and cultural deconstruction fails to stimulate higher-order reasoning (Siregar, 2025). The cognitive impact of HOTS is strictly contingent upon a multidimensional alignment where the curriculum design, instructional delivery, and analytic scoring rubrics form a unified, supportive structure.

### **Contextual Scope and Analytic Limitations**

Acknowledging the perspective of Reviewer 2, the claims regarding HOTS as a pedagogical paradigm must be interpreted within the boundaries of this study's qualitative parameters. Because this research was conducted within a single institutional context in Indonesia involving a purposive sample of 1 lecturer and 5 students, these findings do not seek statistical generalization across all Arabic education departments. Instead, this study offers analytic generalization (Creswell, 2013). By providing a thick description of the institutional and cultural friction points within this specific program, we illuminate cognitive mechanisms, structural bottlenecks, and systemic patterns that are highly transferable to similar tertiary foreign language environments grappling with modern curriculum mandates.

## **CONCLUSION**

This study concludes that Higher-Order Thinking Skills (HOTS) in tertiary Arabic education function as fragmented technical insertions rather than an integrated pedagogical paradigm. Due to institutional path dependency and a philological tradition that privileges morpho-syntactic accuracy, Arabic is treated as a static object of study rather than an active medium for critical thought. While curriculum blueprints mandate higher-order cognitive domains, actual classroom practices and assessments remain bound to lower-order

orientations. However, emerging student engagement during interpretive tasks signals a promising transitional phase toward collaborative meaning-construction.

The study's distinct novelty lies in exposing a culturally embedded epistemic barrier unique to Arabic departments: a deeply rooted tradition of text-deference within Islamic higher education. Unlike general language scholarship that attributes implementation gaps to generic training deficits, this study demonstrates how an institutional baseline of textual replication and memorization actively resists critical deconstruction. Shifting to a true higher-order paradigm thus requires challenging deeply held epistemic beliefs regarding textual authority in non-Western foreign language contexts. Bridging the gap between curricular rhetoric and operational classroom realities requires targeted structural interventions:

Policy Makers and Curriculum Developers: Replace broad declarative mandates with operational guidelines that map concrete linguistic-cognitive scaffolding pathways anchored to functional language thresholds. Department Chairs and Lecturers: Shift professional development from abstract HOTS terminology toward training instructors as cognitive designers who utilize multidimensional analytical rubrics to prevent high-level tasks from degrading into technical grammar grading. Instructional Material Designers: Redesign textbooks by replacing closed-ended structural drills with open-ended inquiry tasks, contrasting reading perspectives, and conceptual mapping activities.

Given its single-institution focus and small qualitative sample, these findings offer analytic framework generalization rather than statistical scope. To build upon these structural and cultural insights, future research should employ longitudinal, cross-institutional mixed-methods designs. Comparative investigations across diverse institutional models—such as state-administered universities versus traditional private seminaries—are essential to test the transferability of this proposed pedagogical paradigm.

## ACKNOWLEDGMENT

Thank you to all authors who have helped and contributed to the completion of this research. We are especially grateful to the editorial team of *Jurnal Ta'lim Al-'Arabiyyah* for their assistance in processing and publishing this paper. We also extend our gratitude to all parties who have helped and contributed to the completion of the research and preparation of this article. However, this research is far from perfect. We hope this research can help develop Arabic language learning media and technology to be more effective and up-to-date..

## AUTHOR CONTRIBUTIONS STATEMENT

[ST] led the study from conception to completion, coordinating all administrative tasks and overseeing the overall publication process. [N] and [SH] contributed by reviewing and validating the data and the manuscript.

## REFERENCES



- Abosalem, Y. (2016). Assessment techniques and students' higher-order thinking skills. *International Journal of Secondary Education*, 4(1), 1–11. <https://doi.org/10.11648/j.ijsedu.20160401.11>
- Alanazi, A. A., Osman, K., & Halim, L. (2024). Effect of scaffolding strategies and guided discovery on higher-order thinking skills in physics education. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(9), Article em2496. <https://doi.org/10.29333/ejmste/14980>
- Arifin, Z., Saputro, S., & Kamari, A. (2025). The effect of inquiry-based learning on students' critical thinking skills in science education: A systematic review and meta-analysis. *Eurasia Journal of Mathematics, Science and Technology Education*, 21(3), Article em2592. <https://doi.org/10.29333/ejmste/15988>
- Baihaqi, A. I., Syammary, N. A., Masrifah, K., Murtadho, N., & Nurhidayati, N. (2025). Development of HOTS-based Arabic language module to improve students' critical thinking skills. *Arabi: Journal of Arabic Studies*, 10(1), 39–52. <https://doi.org/10.24865/ajas.v10i1.908>
- Braun, V., & Clarke, V. (2021). Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research*, 21(1), 37–47. <https://doi.org/10.1002/capr.12360>
- Faber, T. J. E., Dankbaar, M. E. W., van den Broek, W. W., Bruinink, L. J., Hogeveen, M., & van Merriënboer, J. J. G. (2024). Effects of adaptive scaffolding on performance, cognitive load and engagement in game-based learning: A randomized controlled trial. *BMC Medical Education*, 24(1), Article 943. <https://doi.org/10.1186/s12909-024-05698-3>
- Huda, M. M., & Rais, P. (2021). Improving Arabic language learning based on higher order thinking skills (HOTS) in excellent senior high school. *Fenomena*, 20(2), 283–296. <https://doi.org/10.35719/fenomena.v20i2.68>
- Ibrahim, A., & Barnawi, O. Z. (2022). *Higher education in the Arabian Gulf region*. Routledge. <https://doi.org/10.4324/9781003049609>
- Ilgun Dibek, M., Sahin Kursad, M., & Erdogan, T. (2025). Influence of artificial intelligence tools on higher order thinking skills: A meta-analysis. *Interactive Learning Environments*, 33(3), 2216–2238. <https://doi.org/10.1080/10494820.2024.2402028>
- Kamlasi, I., Pratama, H., Wahyuni, S., & Bahri, S. (2025). Exploring higher-order and lower-order oral questions: A revised Bloom's taxonomy of cognitive aspects in undergraduate EFL thesis defense examinations. *Veredas do Direito*, 22(5), Article e223803. <https://doi.org/10.18623/rvd.v22.n5.3803>
- Kania, N., & Kusumah, Y. S. (2025). The measurement of higher-order thinking skills: A systematic literature review. *Malaysian Journal of Learning and Instruction*, 22(1), 97–116. <https://doi.org/10.32890/mjli2025.22.1.6>
- Masrifah, K., & Ahsanuddin, M. (2025). Development of HOTS-based final semester assessment questions in the Al-‘Ashri Arabic textbook for 10th-grade students of Islamic Senior High School. *Arabiyatuna: Jurnal Bahasa Arab*, 9(2), 755–769. <https://doi.org/10.29240/jba.v9i2.13889>
- Mustapha, M. Z., & Abdullah, N. (2024). The effects of flipped classroom teaching method on science students' higher order thinking skills, engagement and satisfaction. *International Journal of Research and Innovation in Social Science*, 8(3s), 3389–3403. <https://doi.org/10.47772/IJRISS.2024.803245S>
- Nugroho, I., Wilujeng, I., Iman, M. S., & Mahesawati, P. N. (2025). Publication trends on higher-order thinking skills and science process skills in elementary schools (1992–2023): A bibliometric analysis of science education. *Al-Bidayah: Jurnal Pendidikan Dasar Islam*, 17(2), 92–119. <https://doi.org/10.14421/al-bidayah.v17i2.12039>

- Putri, A. K., Fahrnunisa, A. H., Evriliani, A. S., Volosa, E., & Nahadi, N. (2024). Transforming chemistry learning through the development of higher order thinking based on revised Bloom's taxonomy: A systematic review. *Hydrogen: Jurnal Kependidikan Kimia*, 12(3), 589–605. <https://doi.org/10.33394/hjkk.v12i3.12452>
- Rahmati, H., Garmabi, B., Ghazvini, H., Tamijani, S. M. S., Sarvandani, M. N., & Rafeiee, R. (2026). Circadian rhythms and human studies: Methodology, challenges, and future directions in research. *Biological Rhythm Research*, 57(2), 71–100. <https://doi.org/10.1080/09291016.2025.2541247>
- Salido, A., Syarif, I., Suparjan, S., Wana, P. R., Sitepu, M. S., & Melisa, R. (2025). Mapping the landscape of critical thinking skills in higher education in the AI era: A bibliometric and systematic literature review. *Journal of Culture and Values in Education*, 8(2), 139–164. <https://doi.org/10.46303/jcve.2025.23>
- Singh, C. K. S., Tao, H., Ong, E. T., et al. (2023). Teachers' self-assessment of and perceptions on higher-order thinking skills practices for teaching writing. *Pegem Journal of Education and Instruction*, 13(3), 337–349. <https://doi.org/10.47750/pegegog.13.03.34>
- Siregar, T. (2025). *Balancing higher- and lower-order thinking skills: An analysis of lecturers' pedagogical practices in higher education*. SSRN. <https://ssrn.com/abstract=5590290>
- Sukmojati, E., Viana, S. A., Djami, M. K., Fauzi, A., Syafirah, D., & Bana, E. O. D. (2025). Critical thinking in English language teaching: A bibliometric analysis. *F1000Research*, 14, Article 1305. <https://doi.org/10.12688/f1000research.171160.1>
- van Nooijen, C. C. A., de Koning, B. B., Bramer, W. M., et al. (2024). A cognitive load theory approach to understanding expert scaffolding of visual problem-solving tasks: A scoping review. *Educational Psychology Review*, 36(1), Article 12. <https://doi.org/10.1007/s10648-024-09848-3>
- Vargas, J. H., Ccama Ojeda, E. C., Coronado Zapata, C. A., Ascoy Flores, K. A., Hernández Vela, J. A., & Damian Espinoza, Y. E. (2024). Analysis of significant learning in higher education: Usefulness of Fink's taxonomy: A systematic review. *Journal of International Crisis and Risk Communication Research*, 7(S7), Article 1341. <https://doi.org/10.63332/joph.v5i6.2393>
- Yang, A., Sulaiman, N. A., & Yacob, N. S. (2025). Enhancing critical thinking skills for higher education students through English reading modules: A systematic review. *Cogent Education*, 12(1), Article 2587466. <https://doi.org/10.1080/2331186X.2025.2587466>
- Yang, Y., & Heong, Y. M. (2024). The significance of developing advanced higher order thinking skills in China's arts education. *Journal of TVET and Technology Review*, 2(1), 28–37. <https://doi.org/10.30880/jttr.2024.02.01.003>
- Yasmar, R., Al Husna, L., Wargadinata, W., & Nurhadi, N. (2023). Exploration HOTS (Higher Order Thinking Skills) in Arabic textbook. *Tanwir Arabiyyah: Arabic As Foreign Language Journal*, 3(2), 97–110.
- Yücel, A. G. (2025). Critical thinking and education: A bibliometric mapping of the research literature (2005-2024). *Participatory Educational Research*, 12(2), Article 23. <https://doi.org/10.17275/per.25.23.12.2>
- Zohar, A., & Dori, Y. J. (2003). Higher order thinking skills and low-achieving students: Are they mutually exclusive? *The Journal of the Learning Sciences*, 12(2), 145–181. [https://doi.org/10.1207/S15327809JLS1202\\_1](https://doi.org/10.1207/S15327809JLS1202_1)