

## The Nature of Truth in Evaluating Arabic Language Learning Through the Epistemological Representation of Language Proficiency

Adibah Ulaya Putri\*, Sofyan Sauri

Universitas Pendidikan Indonesia, Bandung, Indonesia

\*adibahulayap05@upi.edu

### Abstract

Evaluation in Arabic language learning is generally assumed to represent students' language competence, yet this assumption is rarely examined from an epistemological perspective. In practice, there is often a gap between students' test scores and their actual communicative ability. This condition indicates that the current evaluation system does not always reflect students' real language performance. Therefore, this study aims to examine the concept of truth in Arabic language learning evaluation through the perspectives of correspondence theory, coherence theory, and pragmatism. This study employed a qualitative library research method, analyzing primary and secondary sources, including scientific journal articles and academic works on Arabic language evaluation, epistemology, and the philosophy of education. The findings reveal several interconnected problems. From the perspective of correspondence theory, the dominance of grammar-based written tests often yields scores that do not align with students' actual communicative competence, as evidenced by students who achieve high scores yet still struggle with oral Arabic. From the perspective of coherence theory, there is a mismatch between communicative learning objectives and evaluation instruments that remain focused on grammatical and cognitive aspects. Meanwhile, from a pragmatic perspective, the strong emphasis on grades and passing examinations has shifted evaluation from a reflective learning tool into an end in itself, reducing its contribution to the development of meaningful language skills. This study concludes that improving Arabic language evaluation requires not only technical changes but also a philosophical reorientation toward more authentic and meaningful assessment practices.

**Keywords: Arabic Language; Epistemology; Learning Evaluation; Truth**

### Abstrak

*Evaluasi dalam pembelajaran bahasa Arab umumnya dianggap dapat menggambarkan kemampuan berbahasa peserta didik, tetapi anggapan ini jarang dikaji dari sudut pandang epistemologi. Dalam praktiknya, masih ditemukan perbedaan antara nilai tes dan kemampuan komunikasi peserta didik yang sebenarnya. Penelitian ini bertujuan mengkaji konsep kebenaran dalam evaluasi pembelajaran bahasa Arab melalui teori korespondensi, koherensi, dan pragmatisme. Penelitian menggunakan metode kajian pustaka kualitatif dengan menganalisis berbagai artikel ilmiah dan karya akademik yang berkaitan dengan evaluasi bahasa Arab, epistemologi, dan filsafat pendidikan. Hasil penelitian menunjukkan bahwa penggunaan tes tertulis berbasis gramatikal sering menghasilkan nilai yang tidak sesuai dengan kemampuan komunikasi peserta didik. Selain itu, ditemukan ketidaksesuaian antara tujuan pembelajaran komunikatif dan instrumen evaluasi yang masih berfokus pada aspek tata bahasa serta kognitif. Dari perspektif pragmatisme, orientasi pada nilai dan kelulusan membuat evaluasi lebih berfokus pada hasil tes daripada pengembangan kemampuan berbahasa secara nyata. Penelitian ini menyimpulkan bahwa evaluasi pembelajaran bahasa Arab*

*tidak cukup diperbaiki secara teknis, tetapi juga memerlukan perubahan cara pandang terhadap tujuan evaluasi. Sebagai kontribusi, penelitian ini menawarkan kerangka evaluasi autentik berbasis epistemologi yang menekankan kesesuaian antara hasil evaluasi, tujuan pembelajaran, dan kemampuan berbahasa peserta didik dalam praktik nyata.*

**Kata Kunci:** Bahasa Arab; Epistemologi; Evaluasi Pembelajaran; Kebenaran

## Introduction

Evaluation in Arabic language learning is generally understood as an instrument to measure the extent to which learners' language proficiency has developed (Pujiani, 2024). This claim is widely accepted in educational practice. It serves as the basis for pedagogical decision making ranging from grade determination to curriculum policy without being critically questioned regarding its foundational assumptions. However, in practice, a rather fundamental problem arises, whether the scores produced by an evaluation genuinely reflect actual language ability, or merely represent numbers that capture only certain measured aspects.

This question is not simply a technical matter of measurement but touches on a more fundamental issue namely, whether the evaluations carried out can be epistemologically justified. In Arabic language learning practice, evaluation generally remains oriented toward measuring cognitive aspects through written tests, such as vocabulary mastery, nahwu-sharf rules, and text comprehension (Kabalmay et al., 2025). Meanwhile, language ability is in essence complex and encompasses both receptive and productive skills namely *istima'* (listening), *kalam* (speaking), *qira'ah* (reading), and *kitabah* (writing) (Nuranjali et al., 2025).

This imbalance between the form of evaluation and the nature of language ability gives rise to a phenomenon commonly observed in the field: students who achieve high scores yet remain unable to use Arabic communicatively. This condition not only indicates that the evaluation instruments are incomplete but also reveals a discrepancy between the ability intended to be measured and the ability actually being measured. Such problems cannot be adequately analyzed solely from a pedagogical perspective. In the philosophy of science, every knowledge claim, including the claim that evaluation scores represent language ability requires a justifiable epistemological foundation (Vaillo, 2021).

The correspondence theory, for instance, requires that a proposition can only be considered true if it corresponds to the facts it represents (Asiah, Nursalim and Masitoh, 2023). In the context of evaluation, this means a student's score should directly reflect their actual language ability, not merely their performance in a controlled test situation. Furthermore, coherence theory demands logical consistency among propositions within a knowledge system (Harahap et al., 2024). This implies that truth claims in evaluation must be consistent across all components of assessment, from the concept or ability intended to be measured, to the test items constructed, to the interpretation of results.

Meanwhile, pragmatist theory holds that the truth of an evaluation lies in its usefulness and practical impact on learning (Nzioka, Nyangau and Kagama 2024). The meaning that evaluation results should genuinely assist teachers in making sound decisions to develop students' language abilities. In the context of education, it is important to question whether the evaluation scores currently in use satisfy these criteria of truth. Accordingly, these three theories are not merely self contained philosophical constructs but can also function as analytical lenses that directly examine the validity of the meaning attributed to Arabic language evaluation scores. Some prior studies have

explored various dimensions of evaluation in Arabic language learning, ranging from instrument development to the analysis of assessment models. In 2022, Imaduddin et al focused their study on validity and reliability testing of assessment instruments to ensure the accuracy of measurement of Arabic language ability at the elementary level; their findings indicated that 60% of items were valid and 40% were invalid out of ten multiple-choice Arabic test items, with a reliability coefficient of 0.82, classified as very high (Imaduddin, Maulani and Taufik, 2022).

Subakir and Saepurrohman et al analyzed the importance of implementing Higher Order Thinking Skills (HOTS) assessment models across the four language skills, to align evaluation with 21-century critical and creative thinking competencies (Saepurrohman, Yaya and Majid, 2023; Subakir et al., 2025). Akhyar et al on the other hand, employed the Context, Input, Process, Product (CIPP) model to evaluate the overall effectiveness of Arabic language learning programs, demonstrating that the product aspect, including student learning outcomes is frequently influenced by diverse educational backgrounds (Akhyar and Suhardi, 2025).

These studies have made significant contributions in technical and methodological respects. However, upon critical examination, all of them operate under the same underlying assumption, that evaluation scores necessarily reflect students' language ability without questioning that assumption. Not a single study explicitly interrogates the philosophical validity of this representational claim. No research has specifically examined whether such a claim is philosophically sound. In other words, a gap remains at the epistemological level, as existing studies tend to focus on the technical aspects of assessment rather than questioning whether evaluation can genuinely represent language ability in a valid sense. This is the dimension that has largely gone unaddressed in prior research.

Given this gap, the novelty of the present study lies not in the construction of new evaluation instruments but in its perspective. This study does not merely examine how well evaluation measures language ability; it also questions whether evaluation results can genuinely represent language ability epistemologically. This is what distinguishes the present study from previous research, which has predominantly concentrated on technical and psychometric aspects. This study aims to examine the nature of truth in Arabic language learning evaluation through an epistemological review, employing correspondence theory, coherence theory, and pragmatism as analytical frameworks. In practical terms, this study is expected to encourage Arabic language teachers and curriculum developers to adopt a more critical understanding of evaluation, that a sound evaluation is not only statistically valid, but also grounded in a philosophically justifiable basis of truth.

## **Method**

This study employed a qualitative approach within an interpretive paradigm using document analysis as the research method. The data sources consisted of primary data in the form of scientific journal articles, and academic works on Arabic language learning evaluation, epistemology, philosophy of education, and theories of language proficiency, as well as secondary data comprising supporting documents relevant to the study's focus. The data sources were selected purposively based on their thematic relevance, depth of discussion, and connection to the research's epistemological focus. The main instrument of the study was the researcher, who actively selected, read, interpreted, and analyzed the data systematically and reflectively. Data collection was conducted through a documentation study and an in-depth literature review of the selected sources. The data were analyzed qualitatively through several stages, including data reduction, concept

categorization, meaning interpretation, critical analysis of arguments, and the development of a conceptual synthesis regarding the nature of truth in Arabic language learning evaluation.

## **Result and Discussion**

### **1. The Truth of Evaluation from the Perspective of Correspondence Theory**

In the domain of educational epistemology, truth is closely related to the validity of knowledge. Correspondence theory asserts that a proposition can only be considered true when it corresponds to facts or reality (Asiah et al., 2023). In the context of Arabic language learning evaluation, this means that students' scores should accurately reflect their actual language ability, rather than merely their performance on a test (Baruwa and Shutaleva, 2022). Therefore, for an Arabic language learning evaluation to be considered truthful, its results must be capable of accurately capturing the extent to which students have truly mastered the material taught.

Without such correspondence, the scores produced are nothing more than empty numbers that have lost their epistemological meaning. However, various studies indicate that the correspondence between scores and students' actual abilities is rarely achieved. Alamsyah et al., (2025); Fitriani et al., (2025); and Kabalmay et al., (2025) found that Arabic language evaluation continues to be dominated by lower to intermediate order thinking items that predominantly measure grammar, vocabulary, and word forms through written tests (Alamsyah, Fitroh and Baroroh, 2025; Fitriani, Zaman and Imron 2025; Kabalmay et al., 2025).

The problem is that if only grammatical aspects are measured, the resulting scores merely indicate grammatical mastery rather than comprehensive language ability. Consequently, a discrepancy arises between the claimed ability to be measured and the ability actually measured. A commonly observed example in the classroom is a student who can correctly answer multiple-choice questions about *fi'il mudhari'* patterns yet is unable to respond to simple questions in Arabic conversation, a score of 90, in such a case, does not necessarily reflect genuine communicative ability. Meanwhile, that when scores fail to correspond to students' actual abilities, the truthfulness of evaluation results is legitimately open to epistemological questioning (Al-jarf, 2022).

This problem is not merely a technical matter; it also concerns the accuracy of evaluations in representing language ability. From this standpoint, validity in the psychometric sense, which has long been used to assess the quality of instruments in fact only demonstrates that an instrument functions technically, not that it genuinely reflects language ability in its entirety. Imaduddin et al demonstrate that even a statistically valid instrument does not necessarily represent language ability comprehensively (Imaduddin et al., 2022). Accordingly, a statistically valid instrument is not necessarily epistemologically valid. This is the gap that continues to be overlooked in Arabic language learning evaluation.

### **2. The Truth of Evaluation From the Perspective of Coherence Theory**

Beyond correspondence theory, the truth of evaluation can also be examined through coherence theory that is, the alignment and consistency among learning objectives, the abilities intended to be measured, evaluation instruments, and assessment results (Custers, 2019). These principles remind us that the truth of assessment is not a fixed number, but rather the connection between the test and the reality of learning, as well as its coherence within the curriculum framework and educational objectives.

However, various studies indicate that coherence in Arabic language evaluation remains weak. Jannah et al found that while learning objectives genuinely emphasize

communicative ability, the evaluations employed still predominantly measure grammar (Jannah et al., 2024). As a result, a misalignment arises between the learning objectives and the evaluation method used. The evaluation system claims to measure language ability comprehensively, yet what is actually measured constitutes only a small portion of that ability. This problem is also evident in the misalignment among the curriculum, evaluation instruments, and assessment results. The curriculum targets the four language skills in an integrated manner, yet evaluation predominantly measures reading comprehension. Despite this, evaluation results continue to be regarded as representative of overall language ability. From a coherence perspective, this condition indicates that the conclusions drawn are not fully consistent with the basis on which the assessment rests.

Saepurrohman et al and Haniefa attempted to address this problem by implementing HOTS-based assessment. However, raising the cognitive demand of test items does not automatically resolve the coherence problem, as the primary issue lies not only in the difficulty level of the items but also in the continued exclusion of speaking skills from evaluation (Haniefa, 2022; Saepurrohman et al., 2023). Therefore, improving Arabic language evaluation requires more than simply enhancing the quality of written test items; it also necessitates realigning learning objectives, evaluation instruments, and the interpretation of assessment results.

### **3. The Truth of Evaluation from the Perspective of Pragmatist Theory**

Unlike correspondence and coherence theories, pragmatist theory evaluates truth in terms of its utility and practical impact. In learning assessment, evaluation results are considered meaningful only if they can genuinely be used by students to understand their own abilities, by teachers to improve instruction, and by educational institutions as a basis for policy decisions (Nidawati, 2022). From this perspective, it becomes clear that a sound evaluation is not merely one that produces scores, but one that generates real impact on future learning processes.

In Arabic language evaluation, the critical question is whether the scores students receive are genuinely useful for developing their language ability. Akhyar and Suhardi found that student learning outcomes are frequently influenced by prior educational backgrounds, so that the scores obtained do not necessarily reflect the overall success of the learning process (Akhyar and Suhardi, 2025). As a result, evaluation results have not fully functioned as appropriate feedback for instructional improvement. Another recurring problem is the orientation toward "passing the test."

Teachers and students tend to focus on achieving high scores rather than on developing actual language ability. Learning is consequently directed more toward drilling test items and memorizing grammatical rules than toward using Arabic in communicative contexts. From a pragmatic standpoint, evaluation of this kind has limited meaning, as it fails to have an appropriate impact on students' language development. In contemporary educational discourse, truth is no longer regarded as an unassailable entity, but as something contextual and open to examination (Horsthemke, 2019).

This means that truth in evaluation must always remain open to ongoing reflection. This aligns with the concept of the test of truth developed by Roux, an effort to examine truth through the quality of epistemic reasoning, whether based on correspondence with reality or logical consistency within a knowledge system so that an idea may be accepted, revised, or even rejected (Roux, 2025). In the context of Arabic language evaluation, this means that the truthfulness of evaluation results cannot be judged solely by technical procedures, but must also be assessed in terms of the actual impact on the learning process.

#### 4. Critical Synthesis: The Reduction of Evaluative Meaning and Its Epistemological Implications

Based on the analysis of the three theories of truth, the problem in Arabic language learning evaluation lies not only in the technical weaknesses of the instruments, but also in the narrowing of the meaning of evaluation itself. Evaluation, which should represent language ability in its entirety, has instead been predominantly focused on aspects that are technically easy to measure. This narrowing is evident in three respects. First, from a correspondence perspective, the scores obtained do not yet fully reflect actual language ability. Second, from a coherence perspective, learning objectives, evaluation instruments, and assessment results have not been consistently aligned.

Third, from a pragmatic perspective, evaluation results have not yet had a meaningful impact on students' language development. This condition indicates a crisis of representation in Arabic-language evaluation a situation in which the scores produced cannot be fully trusted as an accurate depiction of students' language ability. Rifaie et al affirm that reliability alone is insufficient to guarantee the truthfulness of the evaluation; it must also be accompanied by representational validity that encompasses all aspects of the ability intended to be measured (Rifaie, Hamza and Elfiky, 2021).

As a final implication, Arabic language-learning evaluation needs to be restructured to take into account the epistemological dimensions of truth. Evaluation should no longer function merely as a measurement tool for assigning numbers or labels, but must also serve as a space for reflecting on the learning process itself. Based on the preceding analysis, this study proposes a conceptual contribution: an epistemology-based authentic evaluation framework as an alternative to the evaluation model that has thus far placed greater emphasis on technical aspects. This framework is built upon three core principles derived from the theories of truth.

The first is the correspondence principle, which emphasizes that evaluation must accurately reflect students' actual language ability. For this reason, evaluation cannot rely solely on written tests; it must also incorporate speaking practice assessment so that students' communicative ability can be directly observed. The second is the coherence principle, which demands alignment among learning objectives, evaluation instruments, and assessment results. If the learning objective is communicative language ability, then evaluation must also encompass *kalam* (speaking) and *istima'* (listening) skills, not only *qira'ah* (reading) and grammatical knowledge.

The third is the pragmatist principle, which emphasizes that evaluation results must be useful and serve as a basis for instructional improvement. Students should be able to use them to understand the development of their own abilities, while teachers should be able to draw on them to refine their instructional strategies. These three principles form the foundation for an evaluation system that is not only technically sound but also grounded in a philosophically justifiable basis of truth. Accordingly, Arabic language learning evaluation should not merely serve as an administrative activity but rather provide a more honest representation of students' language ability.

It is therefore essential for educational practitioners to align what is taught, how it is assessed, and how the results are to be meaningfully understood. Only through such an approach can the truth in evaluation more closely reflect the actual nature of students' language ability. Ultimately, the effort to bring scores closer to actual ability is not merely a matter of method but also of how we understand the meaning of evaluation itself. Evaluation should return to its fundamental function as a reflection of ability, rather than serving as a mere numerical formality. With a more critically informed perspective, assessment systems can be continuously developed to become more relevant and meaningful, while more humanely honoring the full range of students' abilities.

## Conclusion

This study finds that the problems in Arabic language learning evaluation lie not only in technical aspects, but also in more fundamental epistemological concerns. From the perspective of correspondence theory, existing evaluations have not yet fully reflected students' actual communicative ability. From a coherence perspective, misalignments persist among learning objectives, evaluation instruments, and the interpretation of results. From a pragmatist perspective, evaluation results have not yet fully functioned as meaningful information for improving the learning process. These findings indicate that improving evaluation cannot be accomplished by addressing technical instrument aspects alone. Arabic language evaluation must also be grounded in a philosophical foundation that questions whether evaluation results genuinely reflect the ability intended to be measured. Evaluation, therefore, needs to shift from a focus on how well an instrument measures, toward the question of whether the ability being measured truly corresponds to the ability claimed to be represented by the evaluation results. This study remains limited to a conceptual review and has not yet been directly tested in the field. Future research is therefore needed to develop and implement an epistemology-based authentic evaluation model in Arabic language learning practice.

## References

- Akhyar, K. N., & Suhardi, I. (2025). Evaluation of The Arabic Language Learning Program based on The Merdeka Curriculum at a Madrasah Tsanawiyah in Makassar. *Lentera Pendidikan: Jurnal Ilmu Tarbiyah dan Keguruan*, 28(1), 40-67.
- Al-Jarf, R. (2022). Grade Inflation In Language And Translation Courses At Saudi Schools And Universities. *British: Journal of Teacher Education and Pedagogy*, 1(2), 8-25.
- Alamsyah, F., Fitroh, R., & Baroroh, R. U. (2025). Innovative Arabic Vocabulary Assessment Instruments to Enhance the Speaking Skills of Grade X Students at MA Science and Technology Al Marzuqiyah Sumedang. *AL-I'ROBY: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban*, 2(2), 82-96.
- Asiah, S., Nursalim, M., & Masitoh, S. (2023). Implications of the Theory of Coherence, Correspondence and Pragmatism of Truth for Education. *Education Achievement: Journal of Science and Research*, 4(3), 87-95.
- Baruwa, I. B., & Shutaleva, A. (2022). Nature Of Knowledge In Philosophy. *Journal of Education, Society and Behavioural Science*, 35(10), 47-59.
- Custers, E. J. F. M. (2019). Theories Of Truth And Teaching Clinical Reasoning And Problem Solving. *Advances in Health Sciences Education*, 24(4), 839-848.
- Fitriani, N., Zaman, B., & Imron, M. (2025). Pengembangan Evaluasi Pembelajaran Bahasa Arab Berbasis Penilaian Autentik Madrasah Tsanawiyah Miftahul Huda Madura. *Tarunaedu: Journal of Education and Learning*, 3(2), 78-89.
- Haniefah, R. (2022). Implementasi Model Penilaian Hots (Higher Order Thinking Skills) Pada Penilaian Empat Keterampilan Berbahasa Arab. *Ta'limi: Journal of Arabic Education and Arabic Studies*, 1(1), 49-71.
- Harahap, I. A., Rifai, A., Fahmi, H. A., & Salminawati, S. (2024). Basic Knowledge and Criteria for Truthfulness From West and Islamic Perspectives. *Jurnal: Jurnal Program Mahasiswa Kreatif*, 8(1), 83-88.
- Horsthemke, K. (2019). Educational Research, Culturally Distinctive Epistemologies And The Decline Of Truth. *European Educational Research Journal*, 18(5), 513-526.

- Imaduddin, M. F., Maulani, H., & Taufik, I. H. (2022). Test the Validity and Reliability of Arabic Learning Questions. *Arabi: Journal of Arabic Studies*, 7(2), 198-207.
- Jannah, S. A., Madi, F. N. B., Rosyid, A., Muhamad, A., & Husin, S. A. (2024). The Implementation of Authentic Assessment in Arabic Language Learning. *Jurnal Al-Maqayis*, 11(1), 69-87.
- Kabalmay, T., Rahmadewi, S., Aziz, A. N. M. A., & Tantawi, I. (2025). Analyzing Cognitive Skill Balance in Arabic Summative Assessments: A Case Based on LOST-MOST-HOST Framework. *At-Tadzkir: Islamic Education Journal*, 4(2), 163-177.
- Nidawati, N. (2022). Keterkaitan Dan Implikasi Pragmatisme Dalam Pendidikan. *Jurnal Mudarrisuna: Media Kajian Pendidikan Agama Islam*, 12(2), 423-444.
- Nuranjali, N., Husnan, A. H., & Nurjannah, N. (2025). Implementation of Total Quality Management in Improving the Quality of Arabic Language Learning at Man 2 Mataram. *International Journal of Applied and Scientific Research*, 3(8), 625-636.
- Nzioka, J. M., Nyangau, D., & Kagema, D. N. (2024). Justification For The Pragmatic Conception of Truth in Scientific Pursuit of Knowledge. *International Journal Of Research And Innovation In Social Science (IJRISS)*, 9(9), 659-670.
- Pujiani, U. (2024). Assesment System for Arabic Language Learning Within the Independent Curriculum. *Kitaba*, 2(3), 165-173.
- Rifaie, N., Hamza, T. M. A. W., & Elfiky, Y. H. (2021). Validity And Reliability Of The Revised Arabic Language Test For 2-4-Year-Old Children: Cross-Sectional Study. *The Egyptian Journal of Otolaryngology*, 37(1), 1-8.
- Roux, R. (2025). Truth And Transformation: The Test Of Truth In The Development Of New Or Revised Ideas, And Related Values And Beliefs. *Journal of Transformative Education*, 23(1), 9-30.
- Saepurrohman, A., Sunarya, Y., & Majid, M. N. (2023). Higher Order Thinking (HOTS) Skills Assessment Model in Arabic Language Skills Learning in Madrasah. *Tadris Al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban*, 2(1), 1-15.
- Subakir, F., Nurfadlilah, S., Fikriyah, F., Aqila, A. N., & Fatimah, S. (2025). Development of Evaluation Model for Arabic Language Skills Based on HOTS (Higher Order Thinking Skills)/Tathwir Namudhaj Taqwim li-Maharat al-Lughah al-'Arabiyyah Qa'im 'ala Maharat al-Tafkir al-'Ulya. *Jurnal Al Bayan: Jurnal Jurusan Pendidikan Bahasa Arab*, 17(2), 384-404.
- Vaillo, G. (2021). Superficiality and Representation: Adding Aesthetics to “Knowledge without Truth”. *Open Philosophy*, 4(1), 36-57.