



ANALISIS URGENSI DAN IMPLEMENTASI EVALUASI PEMBELAJARAN BAHASA ARAB: TINJAUAN KONSEPTUAL SERTA PRAKTIK PADA EMPAT KEMAHIRAN BERBAHASA

**Marsyanti^{1*}, Muhammad Rahmat Alimin², Muhammad Fathin Haqqar³, Muhammad Riswan
 Liling⁴, Ahmad Dani⁵, Herdah⁶**

^{1*,2,3,4,5,6} IAIN Parepare

*Email: marsyantigradia123@gmail.com, muhammadrahmat.alimin@gmail.com, fathinhaqqar002@gmail.com,
riswan01112003@gmail.com, Ahmaddanicondeng@gmail.com, herdah@iainpare.ac.id

DOI: <https://doi.org/10.37081/jipdas.v6i4.5193>

Abstrak

Evaluasi merupakan komponen penting dalam sistem pembelajaran yang berfungsi mengukur ketercapaian tujuan pendidikan. Namun, dalam praktik pembelajaran Bahasa Arab di berbagai lembaga pendidikan, evaluasi masih didominasi oleh pengukuran kemampuan hafalan mufradat dan penguasaan kaidah nahwu-sharaf. Artikel ini bertujuan menganalisis kegagalan evaluasi program pembelajaran Bahasa Arab berbasis hafalan melalui studi kasus pada sebuah madrasah aliyah unggulan yang memiliki capaian nilai ujian tinggi, tetapi peserta didiknya belum mampu berkomunikasi aktif menggunakan Bahasa Arab dalam situasi nyata. Penelitian menggunakan pendekatan analitis-kritis berbasis studi literatur dan analisis kasus. Hasil kajian menunjukkan bahwa sistem evaluasi yang berorientasi pada hafalan menyebabkan terjadinya ketidaksesuaian antara tujuan pembelajaran komunikatif dengan instrumen penilaian yang digunakan. Dampak pedagogis yang muncul meliputi rendahnya kompetensi komunikatif, terbatasnya keterampilan berbicara dan menyimak, serta berkembangnya budaya belajar yang berorientasi pada tes. Evaluasi tradisional gagal membangun maharah lughawiyah karena hanya mengukur kompetensi linguistik secara parsial tanpa memperhatikan penggunaan bahasa dalam konteks autentik. Oleh karena itu, diperlukan model evaluasi autentik yang menilai performa nyata peserta didik melalui tugas komunikatif, wawancara, presentasi, diskusi, dan penilaian proyek. Artikel ini juga menawarkan prototipe rubrik penilaian komunikatif untuk maharah kalam dan istima' sebagai alternatif evaluasi yang lebih sesuai dengan paradigma pembelajaran bahasa modern.

Kata Kunci: Evaluasi Autentik, Pembelajaran Bahasa Arab, Kompetensi Komunikatif, Maharah Kalam, Maharah Istima'.

1. PENDAHULUAN

Bahasa Arab menempati posisi yang unik dalam sistem pendidikan Islam karena berfungsi sekaligus sebagai bahasa agama, bahasa ilmu pengetahuan, dan bahasa komunikasi internasional. Konsekuensinya, pembelajaran Bahasa Arab tidak dapat direduksi hanya pada penguasaan unsur-unsur linguistik seperti kosakata dan tata bahasa, melainkan harus diarahkan pada pembentukan kompetensi komunikatif yang memungkinkan peserta didik menggunakan bahasa secara efektif dalam konteks sosial yang beragam. Perspektif ini sejalan dengan teori kompetensi komunikatif Dell Hymes yang mengkritik pandangan linguistik struktural karena terlalu menekankan aspek gramatikal dan mengabaikan kemampuan menggunakan bahasa sesuai konteks. Dalam perkembangan selanjutnya, Canale dan Swain memperluas konsep tersebut melalui dimensi kompetensi gramatikal, sosiolinguistik, diskursus, dan strategis yang harus berkembang secara terpadu. Dengan demikian, keberhasilan pembelajaran Bahasa Arab tidak cukup diukur melalui penguasaan aturan bahasa, tetapi harus tercermin dalam kemampuan peserta didik membangun makna, menafsirkan pesan, serta berinteraksi secara komunikatif dalam situasi nyata. (Hymes, 1972; Canale & Swain, 1980)



Meskipun paradigma pembelajaran bahasa modern telah mengalami pergeseran menuju pendekatan komunikatif, praktik evaluasi Bahasa Arab di berbagai lembaga pendidikan Islam masih menunjukkan kecenderungan mempertahankan model pengukuran tradisional yang berorientasi pada hafalan. Evaluasi lebih banyak diwujudkan dalam bentuk tes tertulis mengenai arti mufradat, definisi kaidah nahwu-sharaf, identifikasi bentuk i'rab, atau penerapan aturan gramatikal secara mekanis. Model evaluasi semacam ini mencerminkan dominasi paradigma behavioristik yang memandang keberhasilan belajar sebagai akumulasi informasi yang dapat direproduksi kembali saat ujian. Akibatnya, peserta didik terdorong untuk menghafal jawaban yang diprediksi akan muncul dalam tes daripada mengembangkan kemampuan menggunakan bahasa sebagai sarana komunikasi. Situasi tersebut memperlihatkan adanya ketidaksesuaian epistemologis antara tujuan pembelajaran yang menargetkan kompetensi komunikatif dengan instrumen evaluasi yang hanya mengukur kemampuan reproduktif tingkat rendah. (Brown, 2004; Hardiyanti & Enramika, 2024)

Fenomena paradoks antara tingginya nilai ujian dan rendahnya kemampuan komunikasi peserta didik merupakan indikasi kuat terjadinya kegagalan sistem evaluasi. Dalam perspektif teori asesmen pendidikan, validitas suatu instrumen tidak hanya ditentukan oleh ketepatan pengukuran, tetapi juga oleh kesesuaiannya dengan konstruk kompetensi yang hendak diukur. Ketika tujuan pembelajaran menekankan kemampuan berbicara, menyimak, membaca, dan menulis secara komunikatif, sedangkan evaluasi hanya berfokus pada hafalan kosakata dan aturan gramatikal, maka instrumen tersebut mengalami construct underrepresentation atau kegagalan merepresentasikan kompetensi yang sesungguhnya. Akibatnya, skor tinggi yang diperoleh peserta didik tidak mencerminkan kemampuan bahasa yang autentik, melainkan sekadar kemampuan mengingat informasi linguistik dalam jangka pendek. Kondisi ini menjelaskan mengapa banyak peserta didik mampu menjawab soal ujian dengan baik tetapi mengalami kesulitan ketika diminta melakukan percakapan sederhana dalam Bahasa Arab. (Messick, 1995; Fauzi & Rahmawati, 2025).

Kegagalan evaluasi berbasis hafalan juga dapat dianalisis melalui perspektif konstruktivisme sosial yang menempatkan bahasa sebagai alat negosiasi makna dalam interaksi sosial. Dari sudut pandang ini, kemampuan berbahasa tidak berkembang melalui penguasaan aturan secara terisolasi, melainkan melalui pengalaman menggunakan bahasa dalam konteks yang bermakna. Ketika evaluasi hanya mengukur aspek deklaratif berupa pengetahuan tentang bahasa, peserta didik kehilangan kesempatan untuk menunjukkan kemampuan prosedural berupa penggunaan bahasa secara aktual. Sejumlah penelitian mutakhir menunjukkan bahwa pendekatan komunikatif dan aktivitas berbasis tugas mampu meningkatkan keterampilan berbicara, kemampuan berdialog, serta penggunaan kosakata dalam konteks nyata secara lebih signifikan dibandingkan pembelajaran yang berorientasi pada hafalan semata. Temuan tersebut memperkuat argumen bahwa keberhasilan pembelajaran bahasa harus dinilai berdasarkan performa komunikasi, bukan sekadar penguasaan informasi linguistik yang terfragmentasi. (Ekawati, Hidayat, & Rosyada, 2024; Murtadho et al., 2025)

Perkembangan kajian evaluasi pembelajaran bahasa dalam satu dekade terakhir menunjukkan meningkatnya perhatian terhadap asesmen autentik sebagai alternatif terhadap model pengujian tradisional. Asesmen autentik berangkat dari asumsi bahwa kompetensi bahasa hanya dapat diukur secara valid apabila peserta didik dihadapkan pada tugas-tugas yang menyerupai penggunaan bahasa dalam kehidupan nyata. Oleh karena itu, bentuk evaluasi seperti wawancara, presentasi, proyek kolaboratif, portofolio, observasi performa, dan dialog berbasis skenario dinilai lebih mampu merepresentasikan kemampuan komunikatif dibandingkan tes hafalan. Penelitian terbaru dalam pembelajaran Bahasa Arab menunjukkan bahwa asesmen autentik tidak hanya menghasilkan gambaran kemampuan peserta didik yang lebih komprehensif, tetapi juga meningkatkan motivasi belajar, kreativitas, partisipasi aktif, serta kemampuan menggunakan bahasa secara fungsional. Dengan kata lain, asesmen autentik tidak sekadar berfungsi sebagai alat pengukuran, tetapi juga sebagai instrumen pedagogis yang mendorong terjadinya pembelajaran bermakna. (Jannah et al., 2024; Nurjanah et al., 2026; Fauziyah et al., 2025)

Berdasarkan uraian tersebut, kegagalan evaluasi pembelajaran Bahasa Arab berbasis hafalan bukan sekadar persoalan teknis penyusunan instrumen, melainkan mencerminkan persoalan



paradigmatik yang lebih mendasar mengenai cara memaknai kompetensi bahasa. Selama evaluasi masih berorientasi pada reproduksi pengetahuan linguistik, kesenjangan antara nilai akademik dan kemampuan komunikasi akan terus terjadi. Oleh karena itu, diperlukan rekonstruksi sistem evaluasi yang berlandaskan teori kompetensi komunikatif, pendekatan komunikatif, dan prinsip asesmen autentik agar proses pengukuran benar-benar selaras dengan tujuan pembelajaran Bahasa Arab kontemporer. Artikel ini bertujuan mengkaji secara kritis akar kegagalan evaluasi berbasis hafalan dalam pembelajaran Bahasa Arab serta menawarkan model evaluasi autentik yang mampu mengukur dan mengembangkan kompetensi komunikatif peserta didik secara lebih valid, komprehensif, dan kontekstual. (Canale, 1983; Ainin, 2025; Sholahudin et al., 2025).

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif dengan metode analisis kritis berbasis studi kasus. Data diperoleh melalui analisis konseptual terhadap kasus pembelajaran Bahasa Arab pada madrasah aliyah unggulan yang menunjukkan ketidaksesuaian antara nilai akademik dan kemampuan komunikasi peserta didik. Analisis dilakukan melalui tiga tahap, yaitu: 1. Identifikasi kelemahan sistem evaluasi yang digunakan. 2. Analisis dampak pedagogis terhadap kompetensi komunikatif. 3. Perumusan model evaluasi autentik dan rubrik penilaian komunikatif. Kerangka analisis didasarkan pada teori kompetensi komunikatif Hymes, model kompetensi komunikatif Canale dan Swain, serta konsep authentic assessment dalam pembelajaran bahasa.

3. HASIL DAN PEMBAHASAN

A. Kelemahan Sistem Evaluasi yang Digunakan

Evaluasi pembelajaran pada kasus ini menunjukkan dominasi instrumen tes tertulis yang berfokus pada hafalan mufradat dan kaidah nahwu-sharaf. Sistem tersebut memiliki beberapa kelemahan mendasar.

1. Ketidaksesuaian antara Tujuan dan Evaluasi

Kelemahan paling mendasar terletak pada tidak selarasnya tujuan pembelajaran dengan instrumen evaluasi yang digunakan. Pembelajaran Bahasa Arab kontemporer berorientasi pada pengembangan kompetensi komunikatif, tetapi evaluasi hanya mengukur penguasaan kosakata dan kaidah gramatikal. Dalam perspektif validitas konstruk, kondisi ini menunjukkan terjadinya construct underrepresentation, yaitu ketika instrumen gagal merepresentasikan dimensi kompetensi yang menjadi sasaran pembelajaran. Akibatnya, skor yang dihasilkan lebih menggambarkan kemampuan mengingat informasi linguistik daripada kemampuan menggunakan bahasa dalam interaksi nyata. Ketidaksesuaian ini menyebabkan hasil evaluasi kehilangan daya prediktif terhadap performa komunikasi peserta didik di luar kelas. (Canale, 1983)

2. Fokus pada Produk Kognitif Rendah

Dominasi soal hafalan menunjukkan bahwa evaluasi masih beroperasi pada level kognitif rendah dalam taksonomi revisi Bloom, terutama aspek *remembering* dan *understanding*. Padahal kompetensi berbahasa menuntut kemampuan menerapkan, menganalisis, mengevaluasi, bahkan menciptakan tuturan sesuai konteks komunikasi. Penelitian mutakhir menunjukkan bahwa pembelajaran yang hanya mengukur kemampuan reproduksi informasi cenderung menghasilkan pengetahuan inert yang sulit ditransfer ke situasi nyata. Dengan demikian, evaluasi berbasis hafalan tidak hanya membatasi ruang pengembangan kemampuan berpikir tingkat tinggi, tetapi juga menghambat terbentuknya kompetensi komunikatif yang adaptif dan kontekstual. (Ekawati, Hidayat, dan Rosyada, 2024)

3. Mengabdikan Ketrampilan Berbahasa

Evaluasi yang berfokus pada nahwu, sharaf, dan mufradat secara tidak langsung mereduksi hakikat pembelajaran bahasa menjadi sekadar penguasaan sistem linguistik. Padahal teori kompetensi komunikatif menegaskan bahwa kemampuan berbahasa merupakan integrasi antara keterampilan menyimak, berbicara, membaca, dan menulis. Ketika aspek istima' dan kalam tidak diukur secara sistematis, guru kehilangan informasi penting mengenai kemampuan reseptif dan produktif peserta



didik. Dampaknya, keputusan pembelajaran yang diambil berdasarkan hasil evaluasi menjadi kurang akurat karena hanya didasarkan pada sebagian kecil representasi kemampuan bahasa yang sesungguhnya. (Brown, 2004)

4. Tidak Mengukur Penggunaan Bahasa dalam Konteks Nyata

Evaluasi tertulis yang bersifat terisolasi dari konteks komunikasi tidak mampu mengungkap kemampuan peserta didik menggunakan Bahasa Arab secara fungsional. Bahasa pada dasarnya merupakan praktik sosial yang menuntut kemampuan memilih bentuk ujaran sesuai situasi, tujuan, dan lawan bicara. Oleh karena itu, penguasaan kaidah gramatikal tidak secara otomatis menghasilkan kemampuan komunikasi yang efektif. Berbagai penelitian mengenai asesmen autentik menunjukkan bahwa tugas berbasis performa seperti presentasi, dialog, dan wawancara lebih valid dalam mengukur kompetensi bahasa dibandingkan tes hafalan. Dengan demikian, evaluasi yang tidak menghadirkan konteks penggunaan bahasa nyata berpotensi menghasilkan gambaran kemampuan yang semu. (Jannah et al., 2024)

B. Dampak Pedagogis terhadap Kompetensi Komunikatif

Sistem evaluasi yang berorientasi hafalan memberikan dampak signifikan terhadap proses dan hasil pembelajaran.

1. Munculnya Washback Effect Negatif

Teori washback menjelaskan bahwa karakteristik evaluasi akan membentuk perilaku mengajar dan belajar. Ketika evaluasi hanya mengukur hafalan mufradat dan kaidah, guru cenderung mengalokasikan pembelajaran pada latihan soal dan penguasaan materi yang diprediksi keluar dalam ujian. Akibatnya, pembelajaran kehilangan orientasi komunikatif karena aktivitas yang tidak diujikan dianggap kurang penting. Penelitian mutakhir menunjukkan bahwa evaluasi berisiko menghasilkan negative washback ketika instrumen tidak selaras dengan kompetensi yang ditargetkan kurikulum. Dalam konteks Bahasa Arab, kondisi ini mempersempit pengalaman belajar bahasa menjadi sekadar persiapan menghadapi tes. (Katagiri, 2023; Rahman et al., 2023)

2. Rendahnya Kompetensi Sociolinguistik

Kompetensi komunikatif tidak hanya mencakup penguasaan struktur bahasa, tetapi juga kemampuan menggunakan bahasa sesuai norma sosial dan budaya. Evaluasi berbasis hafalan gagal mengukur dimensi sociolinguistik karena peserta didik tidak dituntut mempertimbangkan hubungan penutur, tujuan komunikasi, maupun konteks interaksi. Akibatnya, peserta didik dapat memahami aturan gramatikal namun mengalami kesulitan memilih ungkapan yang tepat dalam situasi nyata. Temuan penelitian terkini menunjukkan bahwa asesmen yang mengabaikan konteks sosial cenderung menghasilkan kompetensi bahasa yang bersifat mekanis dan kurang adaptif terhadap kebutuhan komunikasi autentik. (Sholahudin, Idris, dan Abdurahman, 2025; Jannah et al., 2024)

3. Hambatan pada Maharah Kalam

Perkembangan maharah kalam bergantung pada frekuensi penggunaan bahasa secara aktif, bukan semata-mata penguasaan pengetahuan linguistik. Ketika kemampuan berbicara tidak menjadi bagian dari evaluasi, peserta didik tidak memiliki insentif akademik untuk berlatih berkomunikasi. Situasi ini menciptakan paradoks pembelajaran bahasa, yaitu tingginya pengetahuan tentang bahasa tidak berbanding lurus dengan kemampuan menggunakan bahasa. Berbagai studi eksperimental menunjukkan bahwa peningkatan keterampilan berbicara lebih efektif dicapai melalui tugas performatif, dialog, presentasi, dan interaksi autentik dibandingkan latihan hafalan. Oleh karena itu, absennya asesmen lisan berkontribusi langsung terhadap stagnasi maharah kalam. (Ekawati, Hidayat, dan Rosyada, 2024; Murtadho et al., 2025)

4. Kecemasan Berkomunikasi

Evaluasi yang menekankan jawaban tunggal benar-salah membentuk budaya belajar yang berorientasi pada ketepatan formal. Ketika peserta didik dihadapkan pada komunikasi spontan yang mengandung ketidakpastian, mereka cenderung takut melakukan kesalahan sehingga muncul kecemasan berbahasa (*foreign language anxiety*). Penelitian terbaru pada pembelajar Bahasa Arab menunjukkan bahwa rendahnya kepercayaan diri, pengalaman evaluasi yang menekan, dan minimnya kesempatan berbicara menjadi faktor utama munculnya kecemasan komunikasi. Dengan demikian,



kegagalan evaluasi tidak hanya berdampak pada aspek akademik, tetapi juga membentuk hambatan psikologis yang mengurangi keberanian peserta didik menggunakan Bahasa Arab secara nyata. (Najwa et al., 2024)

5. Terbentuknya Pembelajaran Berorientasi Nilai

Dominasi tes hafalan mendorong terjadinya pergeseran makna keberhasilan belajar dari penguasaan kompetensi menuju pencapaian skor. Dalam perspektif asesmen autentik, kondisi tersebut menunjukkan terjadinya reduksi fungsi evaluasi dari alat peningkatan kualitas pembelajaran menjadi sekadar mekanisme seleksi akademik. Akibatnya, peserta didik mengembangkan strategi belajar jangka pendek yang berfokus pada perolehan nilai, bukan penguasaan keterampilan berbahasa yang berkelanjutan. Sejumlah penelitian terkini menunjukkan bahwa orientasi nilai yang berlebihan berkorelasi dengan rendahnya motivasi intrinsik dan lemahnya transfer kompetensi ke situasi komunikasi nyata. (Ainin, 2025)

C. Mengapa Evaluasi Tradisional Gagal Membangun Maharah Lughawiyah

Kegagalan evaluasi tradisional dapat dijelaskan melalui beberapa perspektif teoritis

1. Perspektif Kompetensi Komunikatif Hymes

Kegagalan evaluasi berbasis hafalan dapat dijelaskan melalui teori kompetensi komunikatif yang dikemukakan Hymes. Menurut perspektif ini, penguasaan bahasa tidak hanya ditentukan oleh ketepatan gramatikal, tetapi juga kemampuan menggunakan bahasa secara tepat sesuai konteks sosial, tujuan komunikasi, dan karakteristik lawan bicara. Evaluasi yang hanya mengukur hafalan mufradat dan kaidah cenderung merepresentasikan bahasa sebagai sistem formal, bukan sebagai praktik sosial. Penelitian terkini dalam pembelajaran Bahasa Arab menunjukkan bahwa dominasi asesmen linguistik menghasilkan kesenjangan antara penguasaan struktur bahasa dan kemampuan komunikasi aktual peserta didik. Dengan demikian, instrumen evaluasi kehilangan relevansi terhadap tujuan penggunaan bahasa yang sesungguhnya. (Sholahudin, Idris, dan Abdurahman, 2025; Jannah et al., 2024)

2. Perspektif Canale dan Swain

Model kompetensi komunikatif Canale dan Swain menegaskan bahwa kemampuan berbahasa merupakan integrasi kompetensi gramatikal, sosiolinguistik, wacana, dan strategis. Evaluasi berbasis hafalan pada umumnya hanya mengukur aspek gramatikal sehingga menghasilkan representasi kompetensi yang parsial. Akibatnya, peserta didik mungkin mampu mengidentifikasi struktur bahasa dengan benar, tetapi mengalami kesulitan membangun koherensi wacana, menyesuaikan tuturan dengan konteks sosial, atau mempertahankan komunikasi ketika menghadapi keterbatasan bahasa. Temuan penelitian terbaru menunjukkan bahwa asesmen yang hanya berorientasi pada tata bahasa berkorelasi dengan rendahnya kemampuan interaksi komunikatif karena tidak memberikan ruang bagi berkembangnya dimensi kompetensi lainnya secara seimbang. (Ainin, 2025; Murtadho et al., 2025)

3. Perspektif Konstruktivisme

Teori konstruktivisme memandang belajar sebagai proses aktif membangun makna melalui interaksi dengan lingkungan dan pengalaman autentik. Dari perspektif ini, evaluasi tidak seharusnya mengukur kemampuan mereproduksi informasi, tetapi kemampuan mengonstruksi pengetahuan untuk menyelesaikan tugas yang bermakna. Evaluasi berbasis hafalan bertentangan dengan prinsip konstruktivistik karena menempatkan peserta didik sebagai penerima informasi pasif yang dinilai berdasarkan ketepatan mengingat materi. Penelitian mutakhir menunjukkan bahwa asesmen autentik berbasis proyek, portofolio, dan performa lebih efektif dalam mendorong pembentukan pemahaman mendalam dibandingkan tes hafalan yang hanya menghasilkan penguasaan pengetahuan jangka pendek dan terfragmentasi. (Fauziyah et al., 2025; Nurjanah, Syawalni, dan Nihayah, 2024)

4. Perspektif Pembelajaran Bahasa Komunikatif

Pendekatan komunikatif menempatkan bahasa sebagai sarana negosiasi makna, bukan sekadar objek kajian linguistik. Oleh karena itu, keberhasilan pembelajaran harus diukur berdasarkan kemampuan peserta didik menggunakan bahasa untuk mencapai tujuan komunikasi tertentu. Ketika evaluasi hanya mengukur hafalan, terjadi ketidaksesuaian antara filosofi pembelajaran dan praktik asesmen sehingga muncul efek washback negatif yang mengarahkan guru dan peserta didik pada aktivitas reproduksi informasi. Berbagai penelitian eksperimental terbaru menunjukkan bahwa



pembelajaran yang didukung asesmen komunikatif menghasilkan peningkatan signifikan pada kemampuan berbicara, interaksi, dan penggunaan bahasa kontekstual dibandingkan evaluasi yang berorientasi pada penguasaan bentuk bahasa semata. (Ekawati, Hidayat, dan Rosyada, 2024; Hardiyanti, Enramika, dan Al Mubarakah, 2024)

D. Model Evaluasi Autentik untuk Maharah Kalam dan Istima'

Sebagai alternatif, diperlukan model evaluasi autentik yang menilai performa nyata peserta didik.

Prinsip Evaluasi Autentik dalam pembelajaran Bahasa Arab.

1. Berbasis Tugas Nyata

Prinsip utama evaluasi autentik adalah penggunaan tugas yang merepresentasikan situasi penggunaan bahasa dalam kehidupan nyata. Berbeda dengan tes hafalan yang mengukur reproduksi pengetahuan, tugas autentik menuntut peserta didik menerapkan kompetensi bahasa untuk menyelesaikan masalah komunikatif yang kontekstual. Dalam perspektif konstruktivisme, pengetahuan memperoleh makna ketika digunakan dalam situasi yang relevan. Penelitian terbaru menunjukkan bahwa asesmen berbasis tugas nyata meningkatkan keterlibatan belajar, transfer kompetensi, dan kemampuan menggunakan Bahasa Arab secara fungsional. Oleh karena itu, validitas evaluasi ditentukan oleh kedekatannya dengan praktik komunikasi yang sesungguhnya. (Fauziyah et al., 2025; Priyandini, 2024)

2. Pengukur Perfoma Aktual

Evaluasi autentik berorientasi pada performa aktual peserta didik dalam menggunakan bahasa, bukan sekadar penguasaan pengetahuan deklaratif tentang bahasa. Prinsip ini berangkat dari asumsi bahwa kompetensi komunikatif hanya dapat diamati melalui tindakan berbahasa yang nyata. Dalam pembelajaran Bahasa Arab, kemampuan berbicara, berdialog, mempresentasikan gagasan, atau memahami pesan lisan lebih representatif dibandingkan kemampuan menghafal kaidah. Penelitian terkini menunjukkan bahwa asesmen performa memberikan gambaran yang lebih komprehensif terhadap kefasihan, koherensi, akurasi, dan kepercayaan diri peserta didik dibandingkan tes konvensional berbasis kertas dan pensil. (Fauzi & Rahmawati, 2025; Fitrawan et al., 2025)

3. Menilai Proses dan Produk

Kelebihan mendasar evaluasi autentik terletak pada kemampuannya menilai hasil akhir sekaligus proses pembelajaran yang melatarbelakanginya. Paradigma asesmen modern memandang bahwa perkembangan kompetensi tidak dapat dipahami hanya melalui skor akhir karena proses belajar mengandung informasi penting mengenai strategi, kreativitas, dan perkembangan peserta didik. Dalam konteks Bahasa Arab, observasi, jurnal refleksi, portofolio, dan proyek memungkinkan guru memperoleh gambaran yang lebih utuh tentang perkembangan kompetensi bahasa. Temuan penelitian menunjukkan bahwa penilaian proses dan produk secara simultan menghasilkan evaluasi yang lebih valid dan mendukung pembelajaran berkelanjutan. (Nurjanah, Syawalni, dan Nihayah, 2025; Jannah et al., 2024)

4. Mengenal berbagai Teknik Penilaian

Kompleksitas kompetensi bahasa menuntut penggunaan berbagai teknik asesmen karena satu instrumen tidak mampu mengukur seluruh dimensi kemampuan berbahasa. Prinsip triangulasi asesmen menekankan pentingnya mengombinasikan observasi, wawancara, presentasi, proyek, portofolio, penilaian diri, dan tes performa. Pendekatan ini memungkinkan guru memperoleh data yang lebih kaya serta mengurangi bias yang muncul akibat ketergantungan pada satu bentuk evaluasi. Penelitian terbaru menunjukkan bahwa diversifikasi teknik asesmen meningkatkan reliabilitas penilaian dan menghasilkan representasi kompetensi peserta didik yang lebih komprehensif dibandingkan evaluasi tunggal berbasis tes tertulis. (Afdal & Sukenti, 2024; Najib & Ridlo, 2024)

5. Berorientasi pada Kompetensi Komunikatif

Orientasi akhir evaluasi autentik adalah mengukur kemampuan peserta didik menggunakan bahasa sebagai sarana komunikasi, bukan sekadar objek pengetahuan. Prinsip ini selaras dengan teori kompetensi komunikatif yang menempatkan kemampuan gramatikal, sosiolinguistik, wacana, dan strategis sebagai satu kesatuan yang tidak dapat dipisahkan. Penelitian mutakhir dalam pembelajaran



Bahasa Arab menunjukkan bahwa asesmen autentik berbasis komunikasi lebih efektif dalam mengembangkan kemampuan interaksi, negosiasi makna, dan penggunaan bahasa secara kontekstual dibandingkan evaluasi yang berfokus pada ketepatan struktur semata. Dengan demikian, evaluasi autentik berfungsi sebagai instrumen pengukuran sekaligus sarana pengembangan kompetensi komunikatif peserta didik. (Sholahudin, Idris, dan Abdurahman, 2025; Mulyani, Habib, dan Hadi, 2025)

Model Evaluasi Maharah Kalam

1. Wawancara Bahasa Arab

Wawancara memungkinkan pengukuran kompetensi komunikatif secara langsung karena peserta didik harus merespons pertanyaan secara spontan. Instrumen ini tidak hanya mengukur akurasi gramatikal, tetapi juga kelancaran, strategi komunikasi, dan kemampuan menyesuaikan tuturan dengan konteks interaksi. (Fitrawan et al., 2025; Murtadho et al., 2025)

2. Presentasi Lisan

Presentasi lisan menuntut integrasi kompetensi wacana, gramatikal, dan strategis dalam penyampaian gagasan secara terstruktur. Penelitian terbaru menunjukkan bahwa asesmen presentasi lebih efektif mengukur kemampuan produksi bahasa dibandingkan tes tertulis karena merefleksikan penggunaan bahasa dalam situasi akademik nyata. (Mulyani, Habib, dan Hadi, 2025; Fauziyah et al., 2025)

3. Role Play

Role play menghadirkan simulasi komunikasi autentik yang memungkinkan peserta didik menggunakan Bahasa Arab sesuai konteks sosial tertentu. Model ini selaras dengan teori kompetensi sosiolinguistik karena keberhasilan komunikasi ditentukan oleh ketepatan penggunaan bahasa, bukan sekadar ketepatan kaidah. (Sholahudin, Idris, dan Abdurahman, 2025; Hardiyanti et al., 2024)

4. Diskusi Kelompok

Diskusi kelompok mengukur kemampuan negosiasi makna, argumentasi, dan kolaborasi linguistik. Dari perspektif konstruktivisme sosial, interaksi antarpeserta didik menjadi sarana pembentukan kompetensi bahasa yang lebih bermakna dibandingkan aktivitas individual berbasis hafalan. (Jannah et al., 2024; Ekawati, Hidayat, dan Rosyada, 2024)

Model Evaluasi Maharah Istima'

1. Menyimak Dialog

Asesmen menyimak dialog memungkinkan pengukuran kemampuan memahami makna eksplisit maupun implisit dalam percakapan. Instrumen ini lebih representatif daripada soal hafalan karena menilai proses pemaknaan bahasa dalam konteks komunikasi yang autentik. (Nurjanah, Syawalni, dan Nihayah, 2025; Fauziyah et al., 2025)

2. Menyimak Video Autentik

Penggunaan video autentik memperluas asesmen istima' melalui integrasi unsur linguistik dan nonlinguistik seperti intonasi, ekspresi, dan konteks budaya. Penelitian mutakhir menunjukkan bahwa media autentik meningkatkan validitas asesmen keterampilan menyimak dibandingkan rekaman yang dibuat khusus untuk tes. (Mulyani, Habib, dan Hadi, 2025; Fitrawan et al., 2025)

3. Dictation (Imla' Masmu')

Imla' masmu' mengukur kemampuan memproses informasi auditori secara akurat sekaligus menghubungkannya dengan representasi tulisan. Model ini memberikan gambaran lebih komprehensif mengenai daya simak, penguasaan kosakata, dan ketepatan identifikasi bunyi bahasa Arab. (Ainin, 2025; Jannah et al., 2024)

4. Information Gap Activity

Information gap activity menempatkan peserta didik pada situasi komunikasi yang mengharuskan pertukaran informasi untuk mencapai tujuan tertentu. Asesmen ini memiliki validitas komunikatif tinggi karena keberhasilan ditentukan oleh efektivitas komunikasi, bukan kemampuan menghafal bentuk bahasa semata. (Murtadho et al., 2025; Ekawati, Hidayat, dan Rosyada, 2024)



E. Prototipe Rubrik Penilaian Komunikatif Bahasa Arab

Rubrik Maharah Kalam

Aspek	Skor 4	Skor 3	Skor 2	Skor 1
Kefasihan	Sangat lancar	Lancar	Kurang lancar	Tidak lancar
Ketepatan Tata Bahasa	Sangat tepat	Cukup tepat	Banyak kesalahan	Dominan salah
Pengucapan	Sangat jelas	Cukup jelas	Kurang jelas	Sulit dipahami
Kosakata	Sangat beragam	Cukup beragam	Terbatas	Sangat terbatas
Interaksi Komunikatif	Aktif merespons	Cukup aktif	Pasif	Tidak mampu merespons

Skor Akhir:

$$\text{Nilai} = (\text{Total Skor} \div 20) \times 100$$

Rubrik Maharah Istima'

Aspek	Skor 4	Skor 3	Skor 2	Skor 1
Pemahaman Isi	Sangat baik	Baik	Cukup	Kurang
Ketepatan Informasi	Sangat tepat	Tepat	Sebagian tepat	Tidak tepat
Kemampuan Menyimpulkan	Sangat baik	Baik	Cukup	Kurang
Respons terhadap Informasi	Sangat relevan	Relevan	Kurang relevan	Tidak relevan

Skor Akhir:

$$\text{Nilai} = (\text{Total Skor} \div 16) \times 100$$

Implikasi bagi Pengembangan Evaluasi Bahasa Arab

Transformasi evaluasi Bahasa Arab perlu dilakukan dengan menggeser paradigma dari assessment of learning menuju assessment for learning. Evaluasi tidak lagi sekadar berfungsi memberikan nilai, tetapi menjadi sarana pengembangan kemampuan komunikasi peserta didik.

Guru Bahasa Arab perlu mengintegrasikan penilaian performa, portofolio, observasi, proyek, dan asesmen autentik ke dalam proses pembelajaran. Dengan demikian, evaluasi dapat menjadi instrumen yang mendorong peserta didik menggunakan Bahasa Arab secara aktif dan bermakna.

4. SIMPULAN

Akibatnya, hasil evaluasi tidak mampu merepresentasikan kemampuan berbahasa peserta didik secara utuh dan menghasilkan kesenjangan antara nilai akademik dengan performa komunikasi nyata.

Dominasi evaluasi tradisional menimbulkan berbagai konsekuensi pedagogis, seperti rendahnya kemampuan berbicara dan menyimak, berkembangnya budaya belajar berorientasi tes, serta terbatasnya kemampuan peserta didik menggunakan Bahasa Arab dalam konteks autentik. Evaluasi yang berfokus pada aspek gramatikal juga mengabaikan dimensi sosiolinguistik, wacana, dan strategi komunikasi yang merupakan unsur penting dalam kompetensi komunikatif.

Oleh karena itu, transformasi menuju evaluasi autentik menjadi kebutuhan yang tidak dapat ditunda. Penerapan wawancara, presentasi, role play, diskusi kelompok, dan tugas menyimak autentik memungkinkan penilaian kompetensi bahasa secara lebih komprehensif. Didukung rubrik komunikatif yang jelas, evaluasi autentik mampu menghasilkan pengukuran yang lebih valid, relevan, dan selaras dengan tuntutan pembelajaran Bahasa Arab abad ke-21 didik.

5. DAFTAR PUSTAKA

- Ainin, Moh. "Evaluation Models and Digital Assessment Innovation in Arabic Language Learning: Towards Authentic and Competency-Based Evaluation." *Mudalla: Proceeding International Conference on Arabic Language* 5, no. 2 (2025). <https://proceeding.arab-um.com/index.php/mudalla/article/view/6>
- Afdal, Adristi, and Desi Sukenti. "Enhancing Authentic Assessment Through Pedagogical Competence: Insights from Indonesian Language Teachers in Riau." *Al-Ishlah: Jurnal*



- Pendidikan* 16, no. 4 (2024): 5127–5141, khususnya 5132–5138. <https://journal.staihubbulwathan.id/index.php/alishlah/article/view/6193>
- Brown, H. Douglas. *Language Assessment: Principles and Classroom Practices*. New York: Pearson Education, 2004.
- Canale, Michael, and Merrill Swain. “Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing.” *Applied Linguistics* 1, no. 1 (1980): 1–47. <https://doi.org/10.1093/applin/1.1.1>
- Canale, Michael. “From Communicative Competence to Communicative Language Pedagogy.” In *Language and Communication*, edited by Jack C. Richards and Richard W. Schmidt, 2–27. London: Longman, 1983.
- Ekawati, Dian, H. D. Hidayat, and Dede Rosyada. “Enhancing Vocabulary, Dialogue, and Writing Skills in Arabic through Communicative Language Teaching: An Experimental Study.” *International Journal of Religion* 5, no. 9 (2024): 896–907. <https://doi.org/10.61707/5880sz57>
- Fauzi, Ahmad, and Siti Rahmawati. “Evaluation of Speaking Competence in Arabic Language Learning: A Performance Assessment Approach.” *International Journal of Arabic Studies* 2, no. 1 (2025): 45–61. <https://ojs.bustanilmu.com/index.php/IJAS/article/view/73>
- Fitrawan, Tomi, R. Umi Baroroh, Oktavia Dwi Lestari, and Khaerul Muttakin. “Authentic Assessment Innovation for Speaking Skills in The Grade 10 Arabic Language Book of The Merdeka Curriculum.” *ARMADA: Jurnal Penelitian Multidisiplin* 4, no. 5 (2025): 421–435, khususnya 425–432. <https://doi.org/10.55681/armada.v4i5.2344>
- Friyandini, Ratna. “Authentic Assessment in Indonesian Language Learning: Theoretical Review and Its Implications.” *Proceeding of International Seminar on Humanity, Education, and Language* 1, no. 1 (2024): 112–124, khususnya 116–121. <https://doi.org/10.21009/ishel.v1i1.57721>
- Fauziyah, Luqi, Dewi Nur Masitho, M. Baihaqi, Nila Sari Syafitri, and Hasanah Dwi Yusuf. “Integration of Authentic Assessment in Arabic Language Learning: Theoretical Perspectives and Non-Test Implications.” *Al Mahāra: Jurnal Pendidikan Bahasa Arab* 11, no. 2 (2025): 215–233, khususnya 220–229. <https://doi.org/10.14421/almahara.2025.0112-03Nurjanah>
- Gepira Enung, Maysi Nabila Syawalni, and Wildatun Nihayah. “The Implementation of Authentic Assessment in Arabic Language Learning at Madrasah Aliyah: A Qualitative Study on Learning Processes and Outcomes.” *Gunung Djati Conference Series* 2026. <https://conferences.uinsgd.ac.id/index.php/gdcs/article/view/3325>
- Hymes, Dell. “On Communicative Competence.” In *Sociolinguistics*, edited by J. B. Pride and Janet Holmes, 269–293. Harmondsworth: Penguin Books, 1972.
- Hardiyanti, Putri, Tomi Enramika, and Zakiyatunnisa Al Mubarakah. “Pendekatan Komunikatif dalam Pengajaran Bahasa Arab.” *Al-Tarqiyah: Jurnal Pendidikan Bahasa Arab* 7, no. 1 (2024). <https://doi.org/10.30631/altarqiyah.v7i1.69>
- Jannah, Siti Alfiaatul, Faisol Nasar Bin Madi, Abdur Rosyid, Ahmad Muhamad, and Saifuddin Ahmad Husin. “The Implementation of Authentic Assessment in Arabic Language Learning.” *Jurnal Al-Maqayis* 11, no. 1 (2024): 69–87. <https://doi.org/10.18592/jams.v11i1.11381>
- Katagiri, Nozomu Wong. “How Sociocultural Factors Mediate Washback Effect: A Study of the Hong Kong Diploma of Secondary Education Examination English Language Paper 4.” *Language Testing in Asia* 13, no. 38 (2023): 1–21, khususnya 3–11. <https://doi.org/10.1186/s40468-023-00250-4>
- Murtadho, Muhammad Alfa Choirul, Mirwan Akhmad Taufiq, Hisbullah Huda, and M. Baihaqi. “Bridging Communication and Competence: The Communicative Approach in Arabic Language Education as a Guide for Advancing Learner Interaction and Proficiency.” *Alsuna: Journal of Arabic and English Language* 8, no. 2 (2025). <https://doi.org/10.31538/alsuna.v8i2.7556>
- Mulyani, Rinda Eka, Ikhwanul Habib, and Nur Hadi. “Authentic Assessment Strategies in Arabic Language Learning in the Digital Era.” *Abjadia International Journal of Education* 10, no. 2 (2025): 338–349, khususnya 340–347. <https://doi.org/10.18860/abj.v10i2.32783>



- Najwa, Hanis, Hazrati Yahaya, Ibrahim Youssef Abdelhamid, and Muhammad Hatta Shafri. "The Level of Student Anxiety in Arabic Speaking Skills." *International Journal of Research and Innovation in Social Science* 8, no. 8 (2024): 2563–2571, khususnya 2565–2569. <https://doi.org/10.47772/IJRISS.2024.8080195>
- Najib, Muhammad, and Ubaid Ridlo. "Analisis Deskriptif Komprehensif tentang Asesmen Otentik." 2024, khususnya 8–15. https://www.researchgate.net/publication/379023305_Analisis_Deskriptif_Komprehensif_tentang_Asesmen_Otentik
- Rahman, Kh. Atikur, et al. "Recognizing Stakeholders and Factors Mediating Washback in Language Testing." *Education Research International* (2023): 1–9, khususnya 2–6. <https://doi.org/10.1155/2023/5548723>
- Sholahudin, Shofwan, Ilham Maulana Ali Idris, and Maman Abdurahman. "The Representation of Communicative Competence in the Arabic Language Learning Outcomes of Indonesia's Merdeka Curriculum." *Aphorisme: Journal of Arabic Language, Literature, and Education* 6, no. 2 (2025). <https://doi.org/10.37680/aphorisme.v6i2.8329>