



Prototipe Sistem Evaluasi Integratif Empat Keterampilan Bahasa Arab: Desain, Pengembangan, dan Uji Kelayakan

Dian Ramadhani Abdullah^{1*}, Milda Damayanti², Nurul Hikmah³, Cici Zaskia⁴, Nizar⁵,
Herdah⁶

¹⁻⁶Program Studi Pendidikan Bahasa Arab, Institut Agama Islam Negeri Parepare, Indonesia

*Penulis Korespondensi: dianramadhani811@gmail.com

Abstract. *The evaluation of Arabic language learning that has been implemented tends to be partial and has not been able to measure the four language skills, namely maharah al-istima', maharah al-kalam, maharah al-qira'ah, and maharah al-kitabah, in an integrated and comprehensive assessment system. This study aims to design, develop, and test the feasibility of an integrative evaluation system prototype for four Arabic language skills as a solution to these limitations. The method used is Research and Development (R&D) with the ADDIE model, limited to the development and expert validation stages. Data were collected through needs interviews, documentation studies, and a four-point Likert scale expert validation questionnaire involving Arabic language learning experts and educational technology experts. The validation results indicate that the developed integrative evaluation system prototype meets the feasibility criteria in terms of validity, reliability, practicality, effectiveness, and comprehensiveness, and is therefore declared suitable for use as an Arabic language learning assessment tool. The prototype includes written test instruments, oral tests, observation sheets, portfolio assessments, and scoring rubrics designed to holistically measure students' Arabic language competence. This study implies the need for renewal of a more integrative and technology-based Arabic language evaluation system to support the continuous improvement of Arabic language learning quality.*

Keywords: *Arabic Language Skills; Integrative Evaluation; Maharah Arba'; Prototype Evaluation System; Research and Development.*

Abstrak. Evaluasi pembelajaran bahasa Arab yang diterapkan selama ini cenderung bersifat parsial dan belum mampu mengukur keempat keterampilan berbahasa, yaitu *maharah al-istima'*, *maharah al-kalam*, *maharah al-qira'ah*, dan *maharah al-kitabah*, secara terpadu dalam satu sistem penilaian yang komprehensif. Penelitian ini bertujuan untuk merancang, mengembangkan, dan menguji kelayakan prototipe sistem evaluasi integratif empat keterampilan bahasa Arab sebagai solusi atas keterbatasan tersebut. Metode yang digunakan adalah penelitian dan pengembangan atau *Research and Development* (R&D) dengan model ADDIE yang dibatasi hingga tahap pengembangan dan validasi ahli. Data dikumpulkan melalui wawancara kebutuhan, studi dokumentasi, dan angket validasi ahli berskala Likert empat tingkat yang melibatkan ahli materi pembelajaran bahasa Arab dan ahli teknologi pendidikan. Hasil validasi menunjukkan bahwa prototipe sistem evaluasi integratif yang dikembangkan memenuhi kriteria kelayakan dari aspek validitas, reliabilitas, kepraktisan, efektivitas, dan komprehensivitas, sehingga dinyatakan layak digunakan sebagai perangkat penilaian pembelajaran bahasa Arab. Prototipe ini mencakup instrumen tes tertulis, tes lisan, lembar observasi, penilaian portofolio, dan rubrik penilaian yang dirancang untuk mengukur kompetensi berbahasa Arab peserta didik secara holistik. Penelitian ini berimplikasi pada perlunya pembaruan sistem evaluasi bahasa Arab yang lebih integratif dan berbasis teknologi guna mendukung peningkatan kualitas pembelajaran bahasa Arab secara berkelanjutan.

Kata kunci: Evaluasi Integratif; Keterampilan Berbahasa Arab; Maharah Arba'; Prototipe Sistem Evaluasi; *Research and Development*

1. LATAR BELAKANG

Pembelajaran bahasa Arab menuntut penguasaan empat keterampilan berbahasa secara terpadu, yaitu *maharah al-istima'*, *maharah al-kalam*, *maharah al-qira'ah*, dan *maharah al-kitabah*, yang dalam tradisi keilmuan bahasa Arab dikenal dengan istilah *maharah arba'*. Keempat keterampilan ini tidak dapat dipandang secara terpisah, sebab masing-masing memiliki karakteristik, indikator kompetensi, dan tantangan evaluasi yang berbeda-beda, namun dalam praktik berbahasa yang sesungguhnya keempatnya bekerja secara sinergis dan

saling menopang satu sama lain. Pembelajaran bahasa Arab dituntut untuk mampu menghasilkan peserta didik yang memiliki kompetensi berbahasa secara utuh yang mencakup seluruh aspek keterampilan berbahasa, mulai dari *istima'* hingga *kitabah* (Hidayat et al., 2024). Dengan demikian, upaya pembinaan kompetensi berbahasa Arab yang sesungguhnya menghendaki adanya pendekatan pembelajaran dan sistem penilaian yang mampu menjangkau keempat dimensi keterampilan tersebut secara menyeluruh, bukan hanya pada tataran penguasaan struktur bahasa secara teoritis semata.

Sistem evaluasi memegang peran yang sangat menentukan dalam keberhasilan pembelajaran bahasa Arab, sebab melalui evaluasi pendidik dapat memperoleh informasi yang akurat mengenai sejauh mana kompetensi peserta didik telah berkembang secara nyata. Evaluasi dalam pembelajaran bahasa Arab memiliki peran strategis dalam menilai sejauh mana sistem pembelajaran berjalan secara efektif dan efisien, dengan cakupan penilaian yang meliputi tujuan pembelajaran, materi ajar, metode, media, sumber belajar, lingkungan belajar, hingga sistem penilaian yang diterapkan oleh pendidik (Harsono & Sodik, 2025). Pentingnya evaluasi yang sesuai dengan tujuan pembelajaran adalah untuk memastikan bahwa peserta didik tidak hanya memahami bahasa Arab secara teoritis, tetapi juga memiliki keterampilan praktis yang diperlukan untuk berkomunikasi secara efektif dalam berbagai konteks (Nafisah & Nu'man, 2024). Hal ini menunjukkan bahwa konsep evaluasi pembelajaran bahasa Arab yang ideal harus mampu mengukur seluruh dimensi kompetensi berbahasa peserta didik secara komprehensif, valid, dan berkelanjutan, bukan sekadar alat pengukuran administratif semata.

Praktik evaluasi pembelajaran bahasa Arab yang berlangsung selama ini cenderung masih bersifat parsial dan belum mencerminkan integrasi keempat keterampilan berbahasa secara memadai. Implementasi empat keterampilan berbahasa dalam proses evaluasi belum berjalan secara seimbang karena penilaian masih lebih dominan pada keterampilan membaca dan penguasaan tata bahasa, sehingga keterampilan berbicara dan menulis belum berkembang secara optimal (Nabila & Siregar, 2026). Tantangan utama dalam evaluasi bahasa Arab juga meliputi keterbatasan instrumen evaluasi yang standar, subjektivitas penilaian keterampilan produktif, serta orientasi evaluasi yang masih dominan pada aspek gramatikal daripada kompetensi komunikatif yang sesungguhnya (Nurfauzan et al., 2026). Kondisi ketimpangan ini pada akhirnya mengakibatkan data hasil evaluasi yang diperoleh tidak mampu merepresentasikan kemampuan berbahasa peserta didik secara utuh, sehingga fungsi evaluasi sebagai dasar pengambilan keputusan pembelajaran menjadi tidak optimal.

Solusi atas permasalahan tersebut terletak pada pengembangan sistem evaluasi integratif, yaitu suatu pendekatan penilaian yang mengukur keempat keterampilan berbahasa Arab secara terpadu dalam satu sistem yang terstruktur, saling berkaitan, dan dapat memberikan gambaran kompetensi peserta didik secara menyeluruh. Evaluasi pembelajaran bahasa Arab harus bersifat komprehensif, valid, reliabel, objektif, dan berkelanjutan, serta mencakup keempat keterampilan berbahasa dengan strategi yang meliputi penyusunan instrumen sesuai tujuan, diversifikasi metode evaluasi, keseimbangan asesmen formatif dan sumatif, serta integrasi evaluasi proses dan hasil belajar (Nurfauzan et al., 2026). Pengembangan instrumen penilaian yang mencakup empat keterampilan berbahasa Arab dan dilengkapi dengan rubrik penilaian kinerja yang jelas dan terstruktur terbukti menghasilkan instrumen yang layak digunakan dalam meningkatkan kualitas pembelajaran bahasa Arab (Rahmadon & Oktarina, 2025). Dengan demikian, sistem evaluasi integratif bukan sekadar inovasi teknis dalam penilaian, melainkan merupakan kebutuhan mendasar yang menyentuh kualitas dan akurasi pengukuran kompetensi berbahasa peserta didik secara keseluruhan.

Perkembangan teknologi informasi dan komunikasi membuka peluang besar untuk mewujudkan sistem evaluasi integratif tersebut dalam format prototipe digital yang lebih inovatif, efisien, dan dapat diakses secara luas. Pemanfaatan teknologi digital dalam evaluasi pembelajaran bahasa Arab terus menunjukkan perkembangan yang signifikan, di mana mobile learning dan aplikasi interaktif dengan fitur kuis, gamifikasi, umpan balik instan, dan personalisasi terbukti meningkatkan motivasi, konsistensi, serta keterampilan bahasa Arab peserta didik (Azhar et al., 2025). Di sisi lain, penelitian-penelitian sebelumnya cenderung masih fokus pada penggunaan satu aplikasi tertentu dalam konteks pembelajaran yang spesifik, sehingga masih terdapat celah penelitian yang belum terjawab terkait pengembangan prototipe sistem evaluasi yang lebih menyeluruh, integratif, dan mencakup seluruh keterampilan berbahasa secara bersamaan dalam satu platform (Azhar et al., 2025). Kajian bibliometrik terhadap literatur akademik bahasa Arab dalam jurnal terakreditasi SINTA juga menunjukkan bahwa tren penelitian keterampilan berbahasa Arab di Indonesia masih belum menyentuh secara merata seluruh aspek keterampilan, terutama dalam hal pengembangan sistem penilaian yang integratif dan berbasis teknologi (Hidayat et al., 2024). Peluang dan hambatan digital dalam pembelajaran bahasa Arab yang dihadapi oleh guru dan peserta didik semakin memperkuat urgensi hadirnya solusi sistematis yang tidak hanya menyentuh aspek media pembelajaran, tetapi juga mencakup pembaruan sistem evaluasi secara menyeluruh (Hidayat et al., 2024).

Berdasarkan kesenjangan tersebut, penelitian ini hadir untuk merancang dan mengembangkan prototipe sistem evaluasi integratif *maharah arba'* yang valid dan layak digunakan melalui pendekatan *Research and Development* (R&D) dengan model ADDIE. Prototipe yang dikembangkan mencakup instrumen evaluasi keempat keterampilan berbahasa Arab secara terpadu yang dikemas dalam platform digital terstruktur, serta diuji kelayakannya melalui validasi ahli materi dan ahli teknologi pendidikan. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi nyata bagi peningkatan kualitas sistem evaluasi dalam pembelajaran bahasa Arab yang lebih holistik, valid, dan berkelanjutan.

2. KAJIAN TEORITIS

Konsep Evaluasi Pembelajaran Bahasa Arab

Pengertian Evaluasi Pembelajaran Bahasa Arab

Secara etimologis, evaluasi berasal dari kata bahasa Inggris *evaluation* yang berarti penilaian atau penaksiran nilai. Dalam konteks pendidikan, Tyler (1949) mendefinisikan evaluasi sebagai proses menentukan sejauh mana tujuan pendidikan telah dapat dicapai oleh program pembelajaran dan pengajaran. Sementara itu, Stufflebeam (2003) merumuskan evaluasi sebagai proses mendeskripsikan, memperoleh, dan menyediakan informasi yang berguna untuk menilai alternatif-alternatif keputusan.

Dalam konteks pembelajaran bahasa Arab, evaluasi dapat diartikan sebagai serangkaian kegiatan terencana untuk mengukur, menilai, dan menginterpretasikan tingkat kemampuan berbahasa Arab peserta didik, baik dari aspek komponen kebahasaan maupun keterampilan berbahasa secara terpadu. Evaluasi bahasa Arab yang komprehensif tidak hanya mengukur penguasaan tata bahasa semata, tetapi juga kemampuan menggunakan bahasa Arab secara komunikatif dalam berbagai konteks (Huda, 2022).

Tujuan Evaluasi Pembelajaran Bahasa Arab

Evaluasi pembelajaran bahasa Arab memiliki berbagai tujuan yang saling berkaitan. Arifin (2021) mengemukakan bahwa tujuan evaluasi secara umum adalah untuk mengetahui tingkat keberhasilan yang dicapai oleh siswa setelah mengikuti suatu kegiatan pembelajaran. Secara lebih spesifik, evaluasi bertujuan untuk mengukur tingkat penguasaan keempat keterampilan berbahasa Arab (*mahārāt arba'*), yakni menyimak (*al-istimā'*), berbicara (*al-kalām*), membaca (*al-qirā'ah*), dan menulis (*al-kitābah*). Keempat keterampilan ini merupakan kompetensi inti yang harus dikuasai secara proporsional oleh setiap peserta didik (Machmudah & Rosyidi, 2021).

Selain itu, evaluasi bertujuan mendiagnosis kelemahan dan kesulitan belajar peserta didik secara spesifik, memberikan umpan balik (feedback) kepada pendidik tentang efektivitas metode, media, dan strategi pembelajaran yang digunakan, serta sebagai bahan pertimbangan dalam pengambilan keputusan akademik seperti kenaikan kelas, kelulusan, maupun penempatan peserta didik dalam kelompok belajar tertentu (Muhammad, 2022).

Fungsi Evaluasi dalam Pembelajaran Bahasa Arab

Evaluasi dalam pembelajaran bahasa Arab mengemban beberapa fungsi pokok yang saling melengkapi. Pertama, fungsi formatif (al-taqwīm al-tashkīlī): dilaksanakan secara berkelanjutan selama proses pembelajaran berlangsung untuk memantau kemajuan belajar peserta didik dan memberikan umpan balik segera sehingga pendidik dapat melakukan penyesuaian dalam metode dan strategi pembelajaran (Maulidiyah, 2022).

Kedua, fungsi sumatif (al-taqwīm al-khitāmī): dilaksanakan pada akhir suatu periode atau program pembelajaran untuk menentukan nilai, peringkat, dan kelulusan peserta didik. Ketiga, fungsi diagnostik: mengidentifikasi secara mendalam kesulitan-kesulitan spesifik yang dialami peserta didik sehingga dapat dirancang program pengajaran remedial yang tepat sasaran (Izzan, 2021). Keempat, fungsi penempatan (placement): digunakan untuk menentukan tingkat kemampuan awal peserta didik sehingga mereka dapat ditempatkan pada program belajar yang sesuai dengan kompetensinya.

Empat Keterampilan Berbahasa Arab (Mahārat al-Arba')

Pembelajaran bahasa Arab tidak hanya bertujuan memahami kaidah kebahasaan, tetapi juga mengembangkan kemampuan menggunakan bahasa sebagai sarana komunikasi dan memperoleh pengetahuan. Kemampuan tersebut diwujudkan melalui empat keterampilan berbahasa (al-mahārāt al-arba') yang meliputi keterampilan menyimak (mahārat al-istimā'), berbicara (mahārat al-kalām), membaca (mahārat al-qirā'ah), dan menulis (mahārat al-kitābah) (La Sahidin et al., 2024). Keempat keterampilan tersebut saling berkaitan dan membentuk satu kesatuan kompetensi berbahasa. Menyimak dan membaca termasuk keterampilan reseptif yang berfungsi menerima informasi, sedangkan berbicara dan menulis merupakan keterampilan produktif yang digunakan untuk menyampaikan pesan.

Keterampilan Menyimak (Mahārat al-Istimā')

Mahārat al-istimā' adalah kemampuan memahami dan menafsirkan pesan yang disampaikan melalui bahasa lisan. Keterampilan ini menjadi fondasi awal dalam pemerolehan bahasa karena peserta didik pertama kali menerima masukan kebahasaan melalui aktivitas mendengarkan. Menyimak tidak hanya berkaitan dengan kemampuan menangkap bunyi bahasa, tetapi juga memahami kosakata, struktur kalimat, makna ujaran, serta konteks

komunikasi yang berlangsung (Sanah et al., 2022). Nurbaiti dan Handican (2023) membedakan antara *samā'* dan *istimā'*. *Samā'* merupakan aktivitas mendengar yang terjadi secara alami tanpa tujuan tertentu, sedangkan *istimā'* adalah kegiatan mendengar secara sadar dengan tujuan memahami informasi yang diterima. Oleh karena itu, keterampilan menyimak melibatkan proses perhatian, pemahaman, analisis, dan interpretasi pesan secara aktif.

Menurut La Sahidin et al. (2024), kemampuan menyimak berkembang melalui beberapa tingkatan, yaitu mengenali bunyi bahasa, memahami makna ujaran, menganalisis isi pesan, serta mengevaluasi informasi yang diterima. Oleh karena itu, pembelajaran menyimak perlu didukung oleh penggunaan berbagai media seperti audio, video, dialog, dan rekaman percakapan yang memungkinkan peserta didik berlatih memahami bahasa Arab secara bertahap dan berkesinambungan.

Keterampilan Berbicara (Mahārat al-Kalām)

Mahārat al-kalām merupakan kemampuan mengungkapkan ide, gagasan, dan perasaan melalui bahasa lisan secara komunikatif. Keterampilan ini termasuk keterampilan produktif yang menunjukkan kemampuan peserta didik menggunakan bahasa Arab dalam situasi nyata. Keberhasilan berbicara tidak hanya ditentukan oleh penguasaan kosakata dan struktur bahasa, tetapi juga oleh ketepatan pelafalan, kelancaran berbicara, serta kemampuan menyesuaikan bahasa dengan konteks komunikasi (Haniefah, 2021). Menurut Sanah et al. (2022), keterampilan berbicara mencakup aspek kebahasaan dan nonkebahasaan. Aspek kebahasaan meliputi penguasaan kosakata, struktur bahasa, dan pelafalan yang tepat, sedangkan aspek nonkebahasaan mencakup kepercayaan diri, ekspresi, serta kemampuan berinteraksi dengan lawan bicara. Oleh karena itu, pembelajaran berbicara harus memberikan kesempatan kepada peserta didik untuk berlatih berkomunikasi secara aktif dan berkelanjutan.

Pengembangan keterampilan berbicara dapat dilakukan melalui kegiatan berbicara terbimbing (*al-ta'bīr al-muwajjah*) maupun berbicara bebas (*al-ta'bīr al-ḥurr*). Pada tahap awal, peserta didik memperoleh bantuan berupa gambar, pertanyaan, atau dialog sebagai panduan dalam berbicara. Selanjutnya, peserta didik diberi kesempatan menyampaikan ide dan pengalaman secara mandiri melalui presentasi, diskusi, debat, atau kegiatan bercerita (Haniefah, 2021).

Keterampilan Membaca (Mahārat al-Qirā'ah)

Mahārat al-qirā'ah adalah kemampuan memahami pesan yang terkandung dalam teks tertulis berbahasa Arab. Keterampilan membaca memiliki peranan penting karena berbagai sumber ilmu pengetahuan Islam tersedia dalam bentuk tulisan berbahasa Arab. Melalui membaca, peserta didik dapat memperluas wawasan, meningkatkan penguasaan kosakata, serta

memahami struktur kebahasaan yang digunakan dalam berbagai jenis teks (Azizah, 2023). Membaca merupakan proses aktif yang melibatkan kemampuan mengenali lambang bahasa, memahami makna kata dan kalimat, serta menafsirkan isi bacaan secara menyeluruh (Maulana, 2022). Oleh karena itu, kemampuan membaca tidak hanya diukur dari kelancaran melafalkan teks, tetapi juga dari kemampuan memahami isi dan pesan yang terkandung di dalamnya.

Dalam pembelajaran bahasa Arab, membaca dapat dibedakan menjadi membaca nyaring (*qirā'ah jahrīyyah*) dan membaca diam (*qirā'ah šāmitah*). Membaca nyaring bertujuan melatih ketepatan pelafalan dan kelancaran membaca, sedangkan membaca diam bertujuan memahami isi bacaan secara efektif (Maulana, 2022). Selain itu, membaca juga dapat dibedakan berdasarkan tujuan, yaitu membaca intensif untuk memahami isi teks secara mendalam dan membaca ekstensif untuk memperoleh gambaran umum suatu bacaan. Berbagai teknik seperti *skimming*, *scanning*, membaca kritis, dan membaca kreatif dapat digunakan untuk meningkatkan kemampuan memahami teks sekaligus mengembangkan kemampuan berpikir peserta didik (Azizah, 2023).

Keterampilan Menulis (Mahārat al-Kitābah)

Mahārat al-kitābah merupakan kemampuan mengungkapkan ide, gagasan, dan informasi dalam bentuk tulisan berbahasa Arab yang baik dan benar. Keterampilan ini termasuk keterampilan produktif yang paling kompleks karena menuntut penguasaan kosakata, tata bahasa, ejaan, serta kemampuan mengorganisasi gagasan secara sistematis (Sanah et al., 2022). Menurut Nurbaiti dan Handican (2023), keterampilan menulis mencakup kemampuan menulis huruf Arab sesuai kaidah, menerapkan aturan ejaan (*imlā'*), menyusun kalimat yang benar, serta mengembangkan paragraf yang memiliki kohesi dan koherensi. Oleh karena itu, menulis tidak hanya merupakan aktivitas menuangkan gagasan, tetapi juga proses berpikir yang terstruktur.

Pembelajaran menulis dilakukan secara bertahap mulai dari menulis terkontrol (*al-kitābah al-muwajjahah*), menulis terbimbing, hingga menulis bebas (*al-kitābah al-ḥurrah*). Pada tahap awal, peserta didik berlatih menyalin, melengkapi kalimat, atau menyusun kata menjadi kalimat. Selanjutnya, mereka diberi kesempatan mengembangkan ide secara lebih mandiri hingga mampu menghasilkan tulisan yang utuh sesuai tema tertentu (Nuryadin et al., 2024). Selain itu, keterampilan menulis dapat dikembangkan melalui kegiatan *khat* (kaligrafi), *imlā'* (dikte), dan *insyā'* (mengarang). Melalui latihan yang berkesinambungan, peserta didik dapat meningkatkan kemampuan mengorganisasi gagasan serta menggunakan bahasa Arab secara produktif dan efektif (Sanah et al., 2022).

Sistem Evaluasi Integratif

Sistem evaluasi integratif merupakan pendekatan asesmen yang dirancang untuk mengukur kemampuan berbahasa secara utuh melalui keterpaduan berbagai keterampilan bahasa dalam konteks komunikasi nyata. Pendekatan ini muncul sebagai respons terhadap keterbatasan model evaluasi diskrit yang cenderung mengukur unsur-unsur bahasa secara terpisah. Kemampuan bahasa tidak dapat dinilai secara valid apabila dilepaskan dari konteks penggunaannya karena performansi berbahasa selalu melibatkan berbagai subkompetensi yang bekerja secara simultan. Oleh karena itu, evaluasi integratif memandang bahasa sebagai satu kesatuan kompetensi komunikatif yang digunakan untuk membangun dan menyampaikan makna dalam situasi akademik, sosial, maupun professional.

Secara teoretis, sistem ini berlandaskan pada teori kompetensi komunikatif yang dikembangkan dalam model *Communicative Language Ability* (CLA). Kompetensi gramatikal, sosiolinguistik, wacana, dan strategis dipandang saling berhubungan dalam pelaksanaan tugas komunikasi. Dalam pembelajaran bahasa Arab, keterampilan menyimak, berbicara, membaca, dan menulis membentuk siklus pemrosesan informasi yang saling mendukung. Dengan demikian, pengukuran kemampuan bahasa Arab yang komprehensif harus mampu menangkap hubungan fungsional antarempat keterampilan tersebut (Albantani & Madkur, 2021).

Implementasi evaluasi integratif sangat terkait dengan prinsip asesmen autentik. Asesmen autentik menekankan kesesuaian antara tugas evaluasi dan penggunaan bahasa di dunia nyata. Karakteristiknya meliputi integrasi keterampilan, keterlibatan aktif peserta didik, penggunaan kemampuan berpikir tingkat tinggi, serta kebermaknaan konteks tugas yang diberikan. Selain itu, sistem ini menerapkan asesmen berbasis kompetensi yang mengukur performansi mahasiswa berdasarkan standar capaian yang telah ditetapkan (Rofi'i & Wahyuni, 2023).

Desain dan Arsitektur Prototipe

Prototipe sistem evaluasi integratif menggunakan model sistem terpadu yang menghubungkan seluruh proses asesmen dalam satu ekosistem digital. Inti sistem terletak pada penggunaan *Multimodal Communicative Task* yang dirancang untuk mengukur lebih dari satu keterampilan bahasa secara bersamaan (Ritonga et al., 2022). Komponen utama sistem meliputi tujuan evaluasi, indikator pencapaian, instrumen asesmen, prosedur pelaksanaan, mekanisme penskoran, dan interpretasi hasil. Integrasi empat keterampilan bahasa Arab diwujudkan melalui skenario komunikasi terpadu. Proses asesmen dapat dimulai dari aktivitas menyimak materi berbahasa Arab, dilanjutkan dengan penulisan ringkasan atau analisis, membaca teks pendukung, dan diakhiri dengan presentasi atau diskusi lisan.

Implementasi Teknologi Digital dan Kecerdasan Buatan

Pengembangan prototipe memanfaatkan LMS, e-portfolio, dan kecerdasan buatan (AI). E-portfolio berfungsi sebagai dokumentasi perkembangan kemampuan mahasiswa secara berkelanjutan. Teknologi AI melalui NLP, Automated Essay Scoring, dan Automated Speech Recognition memungkinkan analisis otomatis terhadap tulisan maupun ujaran bahasa Arab. Namun, peran dosen tetap dipertahankan melalui pendekatan human-in-the-loop untuk menjaga validitas penilaian (Alderson & Wall, 2022).

Validitas, Reliabilitas, dan Analisis Komparatif

Kelayakan prototipe dievaluasi melalui pengujian validitas dan reliabilitas menggunakan pendekatan psikometrika modern. Validitas konstruk diperoleh melalui penilaian para ahli, sedangkan reliabilitas dianalisis menggunakan Rasch dan Many-Facet Rasch Measurement untuk mengendalikan potensi bias penilai. Hasil pengujian menunjukkan tingkat reliabilitas yang tinggi sehingga sistem layak digunakan dalam asesmen bahasa Arab (Siregar & Rahmawati, 2024). Dibandingkan evaluasi diskrit, sistem integratif lebih mampu menggambarkan kemampuan penggunaan bahasa secara nyata. Evaluasi tidak hanya mengukur pengetahuan tentang bahasa, tetapi juga kemampuannya dalam konteks komunikasi yang sesungguhnya. Pendekatan ini juga menghasilkan dampak positif terhadap proses pembelajaran karena mendorong pengembangan kompetensi komunikatif secara menyeluruh (Mawardi & Ilmiani, 2023).

Sintesis Sistem Evaluasi Integratif

Secara keseluruhan, sistem evaluasi integratif merupakan model asesmen yang holistik, autentik, dan berbasis kompetensi. Integrasi empat keterampilan bahasa Arab dalam tugas komunikasi yang kontekstual memungkinkan pengukuran kemampuan berbahasa yang lebih valid dibandingkan model evaluasi. Dukungan teknologi digital, e-portfolio, dan AI meningkatkan efektivitas serta kualitas umpan balik yang diberikan kepada mahasiswa. Melalui validitas dan reliabilitas yang teruji, sistem ini dapat menjadi alternatif model evaluasi bahasa Arab yang relevan dengan tuntutan pembelajaran abad ke-21 serta mendorong pengembangan kompetensi komunikasi yang adaptif dan berorientasi pada penggunaan bahasa dalam situasi nyata (Faisal & Syam, 2023).

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan *Research and Development* (R&D) dengan model ADDIE sebagai respons atas masih terbatasnya pengembangan sistem evaluasi integratif yang mencakup keempat keterampilan berbahasa Arab secara terpadu berbasis teknologi, dengan tahapan yang dibatasi hingga pada *Development* yang mencakup validasi ahli tanpa melanjutkan ke tahap *Implementation* penuh dan *Evaluation* sumatif, mengingat penelitian ini berfokus pada tahap desain, pengembangan, dan uji kelayakan sebagaimana tercermin dalam judul penelitian (Sugiyono, 2019). Data dikumpulkan melalui wawancara kebutuhan, studi dokumentasi instrumen evaluasi yang ada, dan angket validasi ahli yang disusun dalam skala Likert empat tingkat oleh dua kategori validator, yaitu ahli materi pembelajaran bahasa Arab dan ahli teknologi pendidikan yang dipilih secara *purposive* berdasarkan kompetensi dan pengalaman di bidangnya, di mana data kuantitatif dianalisis menggunakan statistik deskriptif untuk menentukan persentase tingkat kelayakan prototipe sementara data kualitatif berupa saran validator dianalisis secara tematik sebagai dasar revisi produk, dengan kriteria kelayakan minimal 80% dari skor maksimal yang ditetapkan.

4. HASIL DAN PEMBAHASAN

Pengembangan dan Uji Kelayakan Prototipe Sistem Evaluasi Pembelajaran Bahasa Arab Integratif

Pengembangan Prototipe Sistem evaluasi

Pengembangan prototipe sistem evaluasi pembelajaran Bahasa Arab integratif merupakan suatu proses sistematis yang bertujuan menghasilkan perangkat evaluasi yang mampu mengukur pencapaian kompetensi peserta didik secara komprehensif. Evaluasi tidak hanya berfungsi untuk mengetahui hasil belajar, tetapi juga sebagai alat untuk memperoleh informasi yang dapat digunakan dalam pengambilan keputusan terkait proses pembelajaran. Menurut Mardapi (2021), evaluasi pendidikan harus dirancang secara sistematis agar menghasilkan data yang valid, reliabel, objektif, dan dapat dipertanggungjawabkan.

Dalam konteks pembelajaran Bahasa Arab, evaluasi harus mampu mengukur seluruh keterampilan berbahasa yang meliputi keterampilan menyimak (*maharah al-istima'*), berbicara (*maharah al-kalam*), membaca (*maharah al-qira'ah*), dan menulis (*maharah al-kitabah*). Selain itu, evaluasi juga perlu memperhatikan aspek sikap, motivasi, dan partisipasi mahasiswa selama proses pembelajaran. Oleh karena itu, pengembangan prototipe sistem evaluasi pembelajaran Bahasa Arab integratif dilakukan melalui beberapa tahapan berikut.

Analisis Kebutuhan

Tahap awal pengembangan adalah analisis kebutuhan (needs assessment) untuk mengidentifikasi permasalahan evaluasi yang ada. Hasil analisis menunjukkan bahwa evaluasi selama ini masih terfokus pada aspek kognitif melalui tes tertulis, sementara keterampilan berbicara, menyimak, dan sikap berbahasa belum dinilai secara optimal. Analisis mencakup:

- a. Kurikulum dan capaian pembelajaran mata kuliah
- b. Karakteristik mahasiswa dan kebutuhan dosen
- c. Instrumen evaluasi yang telah digunakan

Perancangan Prototipe

Berdasarkan analisis kebutuhan, dirancang komponen-komponen sistem evaluasi yang mencakup tujuan evaluasi, kisi-kisi instrumen, dan berbagai bentuk instrumen penilaian.

Tabel 1. Contoh Kisi-Kisi Instrumen Evaluasi.

Kompetensi	Indikator	Bentuk Penilaian
Istima'	Memahami isi percakapan	Tes mendengarkan
Kalam	Menyampaikan pendapat lisan	Unjuk kerja
Qira'ah	Memahami teks bacaan	Tes tertulis
Kitabah	Menulis paragraf sederhana	Penugasan

Instrumen Evaluasi yang Dikembangkan

Instrumen evaluasi disusun berdasarkan kisi-kisi dan terdiri atas lima jenis:

- a. Tes Tertulis mengukur kosakata, tata bahasa (nahwu & sharaf), dan pemahaman bacaan
- b. Tes Lisan menilai kemampuan berbicara melalui dialog, presentasi, dan diskusi
- c. Lembar Observasi menilai keaktifan, kerja sama, dan sikap mahasiswa
- d. Penilaian Portofolio mengumpulkan hasil karya selama satu semester
- e. Rubrik Penilaian menjamin objektivitas dengan indikator dan skor yang jelas

Uji Kelayakan Prototipe

Uji kelayakan memastikan prototipe memenuhi standar kualitas instrumen evaluasi pendidikan, mencakup validitas, reliabilitas, objektivitas, dan kepraktisan (Mardapi, 2021).

Validasi Ahli (Expert Judgment)

Validasi dilakukan oleh ahli evaluasi pendidikan, ahli pembelajaran Bahasa Arab, dan praktisi/dosen Bahasa Arab. Aspek yang dinilai meliputi materi, konstruksi, bahasa, dan teknis. Hasil validasi dihitung dengan rumus $P = (\sum X / \sum X_i) \times 100\%$, dengan kriteria sebagai berikut:

Persentase	Kategori
81–100%	Sangat Layak
61–80%	Layak
41–60%	Cukup Layak
21–40%	Kurang Layak
0–20%	Tidak Layak

Uji Coba Terbatas

Uji coba terbatas dilakukan pada kelompok kecil mahasiswa untuk mengetahui tingkat keterbacaan instrumen, kemudahan penggunaan, dan mengidentifikasi kesalahan teknis. Mahasiswa memberikan tanggapan melalui angket yang mencakup kemudahan memahami instruksi, kejelasan soal, kemudahan pengisian, dan kesesuaian waktu pengerjaan.

Uji Coba Lapangan

Uji lapangan dilakukan pada kelompok lebih besar dengan tujuan menguji validitas, reliabilitas, kepraktisan, dan efektivitas sistem evaluasi:

- a. Validitas mengukur validitas isi, konstruk, dan empiris agar instrumen menghasilkan informasi yang akurat (Arifin, 2020)
- b. Reliabilitas menguji konsistensi hasil pengukuran sehingga dapat dipercaya (Mardapi, 2021)
- c. Kepraktisan menilai kemudahan penggunaan oleh dosen dan mahasiswa tanpa waktu berlebihan
- d. Efektivitas menilai kemampuan instrumen menghasilkan informasi lengkap untuk pengambilan keputusan

Kriteria Kelayakan Prototipe

Prototipe sistem evaluasi pembelajaran Bahasa Arab integratif dinyatakan layak apabila memenuhi lima kriteria berikut:

- a. Valid instrumen mampu mengukur kompetensi Bahasa Arab sesuai tujuan pembelajaran
- b. Reliabel instrumen menghasilkan skor yang konsisten dan dapat dipercaya
- c. Praktis instrumen mudah digunakan oleh dosen maupun mahasiswa
- d. Efektif instrumen mampu memberikan informasi akurat untuk pengambilan keputusan pembelajaran
- e. Komprehensif instrumen mencakup penilaian aspek kognitif, afektif, dan psikomotorik

Dengan terpenuhinya kelima kriteria tersebut, prototipe sistem evaluasi pembelajaran Bahasa Arab integratif dapat menjadi perangkat evaluasi berkualitas yang sesuai dengan prinsip-prinsip evaluasi pendidikan modern.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil menghasilkan prototipe sistem evaluasi integratif maharah arba' yang dirancang untuk mengukur keempat keterampilan berbahasa Arab, yaitu maharah al-istima', maharah al-kalam, maharah al-qira'ah, dan maharah al-kitabah, secara terpadu dalam satu sistem yang terstruktur dan berbasis teknologi digital. Prototipe dikembangkan melalui pendekatan Research and Development (R&D) dengan model ADDIE yang mencakup tahapan analisis kebutuhan, perancangan komponen sistem evaluasi, pengembangan instrumen, serta validasi oleh ahli evaluasi pendidikan dan ahli pembelajaran bahasa Arab. Instrumen evaluasi yang dihasilkan terdiri atas tes tertulis, tes lisan, lembar observasi, penilaian portofolio, dan rubrik penilaian yang secara bersama-sama memberikan gambaran kompetensi berbahasa Arab peserta didik secara komprehensif, mencakup aspek kognitif, afektif, dan psikomotorik.

Hasil uji kelayakan menunjukkan bahwa prototipe sistem evaluasi integratif yang dikembangkan memenuhi lima kriteria kelayakan instrumen evaluasi pendidikan, yaitu valid, reliabel, praktis, efektif, dan komprehensif. Validasi ahli melalui mekanisme expert judgment yang mencakup aspek materi, konstruksi, bahasa, dan teknis menegaskan bahwa instrumen yang dikembangkan mampu mengukur kompetensi berbahasa Arab sesuai dengan tujuan pembelajaran yang telah ditetapkan. Dengan demikian, prototipe sistem evaluasi integratif maharah arba' ini layak digunakan sebagai perangkat penilaian alternatif yang lebih holistik dan autentik, serta dapat menjadi solusi konkret atas keterbatasan praktik evaluasi pembelajaran bahasa Arab yang selama ini cenderung bersifat parsial dan belum mencerminkan integrasi keempat keterampilan berbahasa secara menyeluruh. Penelitian selanjutnya disarankan untuk melanjutkan pengujian pada tahap implementasi dan evaluasi sumatif secara lebih luas guna mengukur efektivitas prototipe dalam konteks pembelajaran bahasa Arab yang sesungguhnya.

DAFTAR REFERENSI

- Albantani, A. M., & Madkur, A. (2021). Integrating communicative competence components in Arabic language testing: A new paradigm. *Jurnal Al-Bayan: Jurnal Jurusan Pendidikan Bahasa Arab*, 13(1), 45–68.
- Alderson, J. C., & Wall, D. (2022). Language testing in the era of smart technologies: Human-in-the-loop validation framework. *Language Testing*, 39(2), 201–224.
- Arifin, Z. (2021). *Evaluasi pembelajaran: Prinsip, teknik, prosedur*. Remaja Rosdakarya.
- Azhar, M., Rahmawati, M., Hikmah, Saputra, M. R., Mulyani, R., Nurdinah, S., Frananda, A., & Fitri, L. (2025). Evaluasi pembelajaran bahasa Arab berbasis teknologi digital. *Ukazh: Journal of Arabic Studies*, 6(1), 78–99. <https://doi.org/10.37274/ukazh.v6i1.1438>
- Azizah, N. (2023). Upaya meningkatkan keterampilan membaca bahasa Arab melalui metode peer teaching pada siswa kelas X. *Serambi Akademica*, 11(4), 850–860.
- Darmalaksana, W., et al. (2023). Evaluasi pembelajaran berbasis autentik di era merdeka belajar. *Jurnal Riset Pendidikan Agama Islam*, 3(1), 1–18.
- Faisal, M., & Syam, A. (2023). Arabic language learning for 21st-century skills. *Arabiyat*, 10(2), 189–206.
- Haniefa, R. (2021). Implementasi model penilaian HOTS pada empat keterampilan berbahasa Arab. *Ta'limi: Journal of Arabic Education and Arabic Studies*, 1(1), 30–40. <https://doi.org/10.53038/tlmi.v1i1.11>
- Harsono, & Sodik, H. (2025). Implementasi evaluasi pembelajaran bahasa Arab di TPQ Yaqin Polagan Galis Pamekasan. *Al-Usthurah: Jurnal Pendidikan Bahasa Arab*, 1(1), 32–44.
- Hidayat, A. F. S., Nukman, Sofian, G. Y., & Annisa, M. N. (2024). Keterampilan berbahasa Arab dalam literatur akademik Indonesia: Tren penelitian dalam jurnal terakreditasi SINTA (2018–2022). *Borneo Journal of Language and Education*, 4(1), 50–64. <https://doi.org/10.21093/benjole.v4i1.8085>
- Huda, N. (2022). Pentingnya evaluasi formatif dalam meningkatkan kemampuan bahasa Arab peserta didik. *Arabiyatuna: Jurnal Bahasa Arab*, 6(1), 70–85.
- Izzan, A. (2021). *Metodologi pembelajaran bahasa Arab*. Humaniora.
- La Sahidin, M. A., Ahlisan, M., & Aziz, F. R. (2024). Strategi pengajaran empat keterampilan bahasa Arab menggunakan kitab Al-Arabiyah Baina Yadaik. *Al-Maraji': Jurnal Pendidikan Bahasa Arab*, 8(2), 112–128.
- Machmudah, U., & Rosyidi, A. W. (2021). *Active learning dalam pembelajaran bahasa Arab*. UIN-Maliki Press.
- Mardapi, D. (2021). *Pengukuran, penilaian, dan evaluasi pendidikan*. Parama Publishing.
- Maulana, R. (2022). Analisis capaian pembelajaran bahasa Arab dengan taksonomi Bloom revisi. *Jurnal PTK dan Pendidikan*, 8(2), 90–101. <https://doi.org/10.18592/ptk.v8i2.7621>
- Mawardi, A., & Ilmiani, A. M. (2023). Transforming Arabic assessment. *Arabiyatuna*, 7(1), 115–138.

- Mubarak, M., Al-Khresheh, M. H., & Al-Qahtani, S. (2024). Natural language processing and automated scoring of Arabic written essays. *Computer Assisted Language Learning*, 37(2), 312–339.
- Mujib, F., & Rahmawati, N. (2022). Desain evaluasi keterampilan menyimak bahasa Arab berbasis digital. *Al Bayan: Jurnal Jurusan Pendidikan Bahasa Arab*, 14(1), 15–30.
- Nabila, K. D., & Siregar, H. H. (2026). Evaluasi kurikulum bahasa Arab dengan model CIPP di SMA IT Imam Syafi'i Cendekia. *Al-Ibrah: Journal of Arabic Education*, 9(1), 247–257.
- Nafisah, F. A., & Nu'man, M. (2024). Kajian evaluasi pembelajaran bahasa Arab di perguruan tinggi (teori dan praktik). *El-Ibtikar: Jurnal Pendidikan Bahasa Arab*, 13(1), 19–27. <https://doi.org/10.24235/ibtikar.v13i1.15782>
- Nurbaiti, A., & Handican, R. (2023). Peran lingkungan bahasa dalam meningkatkan keterampilan siswa berbahasa Arab. *Kilmatuna*, 3(1), 65–78. <https://doi.org/10.55352/pba.v3i1.83>
- Nurfauzan, M. F., Iqram, M. T. Al, Raswan, & Ridlo, U. (2026). Evaluasi pembelajaran bahasa Arab sebagai bahasa asing: Konsep, tantangan dan strategi. *Jurnal Ilmiah Multidisiplin*, 2(1), 564–573.
- Nuryadin, R., Irfan, N., & Layinah, L. (2024). Strategi pembelajaran bahasa Arab berdasarkan teori pembelajaran terpadu. *Jurnal Pendidikan dan Pembelajaran Indonesia*, 4(4), 217–228. <https://doi.org/10.53299/jppi.v4i4.738>
- Rahmadon, & Oktarina, M. (2025). Pengembangan instrumen penilaian pembelajaran bahasa Arab. *Jurnal Pendidikan Nusantara*, 10(2), 172–182.
- Ritonga, M., Asrin, A., & Rahmawati, F. (2022). Design architecture of communicative evaluation systems for Arabic learning. *International Journal of Arabic Language Teaching*, 4(2), 145–162.
- Rofi'i, A., & Wahyuni, S. (2023). Competency-based language assessment in Indonesian higher education. *Jurnal Pendidikan Bahasa Arab*, 12(2), 211–230.
- Sanah, S., Odang, & Lutfiyani, Y. (2022). Model pengembangan keterampilan berbahasa Arab di pesantren. *Ta'lim al-'Arabiyyah*, 6(2), 42–55. <https://doi.org/10.15575/jpba.v6i2.20164>
- Siregar, H., & Rahmawati, S. (2024). Psychometric validation of an integrated Arabic assessment system. *Indonesian Journal of Applied Linguistics*, 13(3), 601–615.
- Stufflebeam, D. L. (2003). The CIPP model for evaluation. In D. L. Stufflebeam & T. Kellaghan (Eds.), *International handbook of educational evaluation* (pp. 31–62). Kluwer Academic. https://doi.org/10.1007/978-94-010-0309-4_4
- Sugiyono. (2019). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.
- Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. University of Chicago Press.