

Evaluasi Tes Kemampuan Bahasa Arab di UIN Syarif Hidayatullah Jakarta: Analisis Kesesuaian dengan Standar CEFR dan ACTL

Ainur Rasidah¹, Khaerul Umam², Fatmah Zaenal Arifin Hambali³, Ubaid Ridho⁴,
Raswan⁵

Program Studi Magister Pendidikan Bahasa Arab Fakultas Tarbiyah Dan Keguruan
Universitas Islam Syarif Hidayatullah^{1,2,3,4,5}

*Email Korespondensi: arrummanu@gmail.com

Sejarah Artikel:

Diterima	Tgl-Bln-Thn
Disetujui	Tgl-Bln-Thn
Diterbitkan	Tgl-Bln-Thn

ABSTRACT

This study aims to analyze the system and implementation of the Test of Arabic as a Foreign Language (TOAFL) at UIN Syarif Hidayatullah Jakarta, and assess its compliance with the international standards of the Common European Framework of Reference for Languages (CEFR) and the American Council on the Teaching of Foreign Languages (ACTFL). The results show that the TOAFL system has a well-established institutional structure and administrative mechanism, but conceptually it is still oriented towards the language knowledge testing paradigm and does not yet measure communicative competence as required by the CEFR and ACTFL. In terms of content, TOAFL test items are dominated by linguistic aspects (grammar and vocabulary), with a very limited portion of pragmatic and socio-cultural abilities. Therefore, TOAFL results cannot be directly converted to CEFR competency levels (A1–C2) or ACTFL categories (Novice–Superior). This study recommends developing an alignment framework between TOAFL and international standards through the development of performative indicators, strengthening oral and written test rubrics, and providing assessment literacy training for Arabic language teachers. These findings are expected to strengthen the competitiveness of Indonesia's Arabic language assessment system in the context of the globalization of Islamic higher education.

Keywords: TOAFL, CEFR, ACTFL, Arabic language assessment, communicative language testing

ABSTRAK

Penelitian ini bertujuan untuk menganalisis sistem dan pelaksanaan Test of Arabic as a Foreign Language (TOAFL) di UIN Syarif Hidayatullah Jakarta, serta menilai kesesuaiannya dengan standar internasional Common European Framework of Reference for Languages (CEFR) dan American Council on the Teaching of Foreign Languages (ACTFL). Hasil penelitian menunjukkan bahwa sistem TOAFL telah memiliki struktur kelembagaan dan mekanisme administrasi yang mapan, namun secara konseptual masih berorientasi pada paradigma language knowledge testing dan belum mengukur kompetensi komunikasi sebagaimana yang dituntut oleh CEFR dan ACTFL. Dari segi konten, butir soal TOAFL didominasi oleh pengukuran aspek linguistik (grammar dan vocabulary), dengan porsi yang sangat terbatas pada kemampuan pragmatik dan sosial-budaya. Oleh karena itu, hasil TOAFL belum dapat dikonversi langsung ke level kompetensi CEFR (A1–C2) atau kategori ACTFL (Novice–Superior). Penelitian ini merekomendasikan pengembangan alignment framework antara

TOAFL dan standar internasional melalui penyusunan indikator performatif, penguatan rubrik tes lisan dan tulisan, serta pelatihan assessment literacy bagi pengajar bahasa Arab. Temuan ini diharapkan dapat memperkuat daya saing sistem evaluasi bahasa Arab di Indonesia dalam konteks globalisasi pendidikan tinggi Islam.

Kata kunci: TOAFL, CEFR, ACTFL, evaluasi bahasa Arab, communicative language testing

Bagaimana Cara Sitasi Artikel ini:

Ainur Rasidah, Khaerul Umam, Fatmah Zaenal Arifin Hambali, Ubaid Ridho, & Raswan. (2025). Evaluasi Tes Kemampuan Bahasa Arab di UIN Syarif Hidayatullah Jakarta: Analisis Kesesuaian dengan Standar CEFR dan ACTL. *Jejak Digital: Jurnal Ilmiah Multidisiplin*, 2(1), 534-546. <https://doi.org/10.63822/86vwfz83>

PENDAHULUAN

Dalam konteks pendidikan tinggi Islam di Indonesia, penguasaan bahasa Arab menjadi elemen esensial yang menopang keberhasilan akademik dan profesional mahasiswa. Bahasa Arab tidak hanya berfungsi sebagai alat komunikasi ilmiah dalam studi keislaman, tetapi juga sebagai sarana akses terhadap khazanah klasik dan modern yang menjadi dasar pengembangan ilmu-ilmu keislaman. Oleh karena itu, evaluasi kemampuan bahasa Arab memiliki posisi strategis dalam memastikan capaian kompetensi linguistik mahasiswa sesuai dengan tujuan pembelajaran yang diharapkan.

Salah satu instrumen yang paling banyak digunakan untuk mengukur kemampuan bahasa Arab mahasiswa di perguruan tinggi Islam adalah TOAFL (Test of Arabic as a Foreign Language). UIN Syarif Hidayatullah Jakarta merupakan pelopor dalam pengembangan dan pelaksanaan TOAFL di Indonesia, menjadikannya model rujukan nasional bagi lembaga-lembaga serupa. Tes ini berperan penting sebagai alat seleksi akademik, prasyarat kelulusan, maupun penentu beasiswa. Namun demikian, dalam era globalisasi pendidikan bahasa, muncul tuntutan baru untuk menyesuaikan sistem evaluasi lokal dengan standar pengukuran kompetensi internasional yang lebih terukur dan diakui secara global.

Perkembangan paradigma evaluasi bahasa modern menekankan kompetensi komunikatif (communicative competence) yang mencakup aspek linguistik, sociolinguistik, dan pragmatik. Standar internasional seperti CEFR (Common European Framework of Reference for Languages) dan ACTFL (American Council on the Teaching of Foreign Languages) telah menjadi acuan global dalam mendesain, menilai, dan menafsirkan kemampuan bahasa asing secara holistik dan berbasis performa. CEFR, misalnya, memetakan kemampuan bahasa dalam enam level (A1–C2) yang menggambarkan sejauh mana pengguna bahasa mampu berinteraksi secara efektif dalam konteks kehidupan nyata. ACTFL, di sisi lain, menilai kemampuan bahasa berdasarkan performa komunikatif dari tingkat *Novice* hingga *Superior* dengan fokus pada akurasi, kelancaran, dan ketepatan fungsi sosial.

Dalam konteks ini, penting untuk mengkaji sejauh mana TOAFL di UIN Syarif Hidayatullah Jakarta telah mencerminkan prinsip-prinsip evaluasi bahasa modern sebagaimana dirumuskan dalam CEFR dan ACTFL. Kajian ini menjadi relevan karena sebagian besar instrumen TOAFL masih menitikberatkan pada aspek struktural dan gramatikal (grammar-based testing), sementara pendekatan global kini mengarah pada penilaian berbasis tugas dan komunikasi nyata (*task-based assessment*). Hal ini menimbulkan pertanyaan mengenai validitas konstruk dan reliabilitas hasil tes, serta sejauh mana hasil TOAFL dapat dibandingkan dengan sistem pengukuran kompetensi bahasa internasional.

Beberapa penelitian sebelumnya di Indonesia menyoroti validitas dan reliabilitas TOAFL dalam konteks nasional, namun belum banyak yang mengaitkannya dengan kerangka internasional seperti CEFR dan ACTFL. Sementara di tingkat global, kajian *alignment* antara tes bahasa seperti TOEFL, IELTS, dan CEFR telah banyak dilakukan, menunjukkan pentingnya harmonisasi antara tes lokal dan standar global untuk menjamin portabilitas hasil dan kredibilitas akademik. Dengan demikian, terdapat kesenjangan penelitian (research gap) yang perlu diisi, yaitu analisis komprehensif terhadap kesesuaian TOAFL sebagai instrumen pengukuran kemampuan bahasa Arab dengan standar kompetensi bahasa yang diakui secara internasional.

Selain itu, fenomena globalisasi akademik menuntut lembaga pendidikan Islam di Indonesia, termasuk UIN Jakarta, untuk menyesuaikan sistem evaluasi bahasanya agar selaras dengan praktik internasional. Hal ini bukan sekadar untuk memenuhi tuntutan administratif, tetapi juga untuk meningkatkan mutu pendidikan bahasa Arab, memperkuat daya saing lulusan di tingkat global, serta memastikan keadilan dan keterukuran hasil evaluasi.

Berdasarkan konteks tersebut, penelitian ini dilakukan untuk menganalisis kesesuaian sistem TOAFL di UIN Syarif Hidayatullah Jakarta dengan standar CEFR dan ACTFL, dengan pendekatan kualitatif deskriptif melalui studi kasus dan analisis isi. Penelitian ini diharapkan tidak hanya memberikan gambaran empiris tentang praktik evaluasi bahasa Arab di perguruan tinggi Islam, tetapi juga menawarkan rekomendasi konseptual untuk pengembangan sistem tes yang lebih valid, reliabel, dan berorientasi global.

Memasuki era modern dan kolonialisme, ilmu pendidikan Bahasa Arab menghadapi tantangan baru. Pengaruh sistem pendidikan Barat dan munculnya bahasa-bahasa kolonial (seperti Inggris, Prancis, dan Belanda) menimbulkan perubahan orientasi dan metodologi pengajaran. Di berbagai wilayah dunia Arab, reformis seperti Rifa'ah al-Tahtāwī dan Muhammad 'Abduh menyerukan perlunya pembaruan pendidikan bahasa agar lebih adaptif terhadap tuntutan modernitas tanpa meninggalkan nilai-nilai Islam. Proses modernisasi tersebut mendorong munculnya disiplin *ta'lim al-lughah al-'Arabiyyah* sebagai bidang kajian akademik tersendiri di berbagai universitas Islam dan lembaga pendidikan tinggi.

Dalam konteks kontemporer (2020–2025), ilmu pendidikan Bahasa Arab mengalami transformasi metodologis dan epistemologis yang signifikan. Paradigma pengajaran beralih dari pendekatan tradisional yang berbasis hafalan dan gramatikal menuju pendekatan komunikatif (*communicative approach*), *task-based learning*, serta integrasi teknologi digital seperti *e-learning*, *multimedia*, dan kecerdasan buatan dalam pembelajaran. Studi-studi terbaru juga menyoroti pentingnya penyesuaian kurikulum dengan kerangka kompetensi global (seperti *Common European Framework of Reference for Languages* – CEFR) agar pembelajaran bahasa Arab memiliki standar internasional dan relevansi praktis.

Di Indonesia, pengajaran Bahasa Arab memiliki sejarah panjang yang berakar pada lembaga keagamaan tradisional seperti pesantren dan madrasah. Namun, hingga kini masih terdapat kesenjangan antara pendekatan tradisional dan kebutuhan pembelajaran modern. Tantangan utama yang dihadapi adalah kurangnya kompetensi komunikatif peserta didik, keterbatasan media pembelajaran, serta belum optimalnya implementasi kurikulum berbasis kompetensi. Oleh karena itu, studi historis terhadap perkembangan ilmu pendidikan Bahasa Arab menjadi penting untuk memahami akar permasalahan dan menemukan solusi inovatif yang kontekstual bagi pengajaran Bahasa Arab di masa kini.

Secara paralel, ACTFL terus memperbarui pedoman profisiensi untuk meningkatkan keterkaitan antara deskripsi performa dan praktik penilaian, dengan revisi signifikan terbaru yang dirilis pada 2024. Revisi ini menegaskan pendekatan performatif yang menilai kemampuan bahasa berdasarkan pencapaian tugas komunikatif yang dapat diamati pada domain mendengarkan, berbicara, membaca, dan menulis. Pembaruan ACTFL menyediakan kerangka operasional yang dapat dimanfaatkan sebagai landasan penilaian rater, skala skor, dan pengembangan rubrik, sehingga kajian kesesuaian TOAFL juga perlu menguji keselarasan kriteria penilaian internal TOAFL terhadap kriteria FACT (Function, Accuracy, Content, Text type) dan deskriptor performa ACTFL.

Dari perspektif metodologis-assessment, literatur kontemporer menekankan dua tuntutan utama: pertama, kebutuhan validasi konstruk yang kuat untuk memastikan bahwa item dan tugas tes memang mengukur konstruk yang dimaksudkan; kedua, literasi assessment dari pengembang dan penilai tes agar praktik scoring konsisten dan adil. Sumber-sumber kajian assessment dan buku rujukan terbaru menunjukkan bahwa tanpa dokumentasi blueprint tes yang transparan dan proses kalibrasi rater yang sistematis, klaim reliabilitas dan validitas menjadi rapuh. Implikasi praktisnya adalah bahwa lembaga yang menyelenggarakan tes seperti TOAFL perlu menunjukkan bukti empiris (analisis item, studi kesetaraan rater, dan uji keterbandingan) sebelum hasil tes dapat diklaim setara atau kompatibel dengan standar internasional.

Kajian awal tentang TOAFL dan praktik penilaian bahasa Arab di Indonesia relatif terbatas dan sebagian besar bersifat deskriptif atau fokus pada reliabilitas internal; studi komparatif yang secara eksplisit memetakan indikator TOAFL ke CEFR atau ACTFL masih jarang. Di samping itu, dokumentasi operasional dari pusat bahasa menunjukkan adanya variasi tujuan tes (mis. seleksi administratif, syarat kelulusan, asesmen pembelajaran), yang berpotensi menimbulkan konflik tujuan antara penggunaan tes untuk keputusan administratif dan untuk penilaian pembelajaran berbasis kompetensi. Kesenjangan penelitian ini—yaitu minimnya analisis alignment antara TOAFL dan kerangka internasional yang terstandar—mendorong kebutuhan penelitian yang menggabungkan analisis isi instrumen, wawancara pengelola, dan persepsi peserta sebagai sumber triangulasi.

Secara sosial-akademik, ketidakselarasan antara tes lokal dan standar internasional memiliki implikasi nyata: hasil tes yang tidak dapat dipetakan ke standar internasional mengurangi portabilitas sertifikat, melemahkan daya saing lulusan pada kancah internasional, serta berpotensi menimbulkan ketidakadilan bagi peserta yang dipengaruhi oleh desain tes yang tidak sesuai tujuan pembelajaran. Dari perspektif kurikulum, ketidaksesuaian antara penilaian dan standar kompetensi global juga menuntut revisi blueprint pembelajaran sehingga proses pembelajaran mendukung pencapaian descriptor-level yang diharapkan. Oleh karena itu, evaluasi kesesuaian TOAFL bukan sekadar exercise teknis, melainkan langkah strategis untuk memperkuat mutu pendidikan bahasa Arab di perguruan tinggi dan menjamin relevansi keluaran pendidikan dengan kebutuhan global.

Penelitian ini menggunakan pendekatan kualitatif deskriptif dengan analisis isi dan studi kasus karena karakter masalah yang menuntut pemahaman mendalam terhadap praktik, dokumen, dan persepsi aktor. Pendekatan ini memungkinkan peneliti menelusuri alasan di balik pilihan desain tes, menilai kecocokan indikator TOAFL dengan deskriptor CEFR dan ACTFL melalui pemetaan rubrik, serta menggali perspektif pihak pengelola dan peserta mengenai validitas, reliabilitas, dan dampak tes terhadap pembelajaran. Pendekatan triangulatif memperkuat validitas temuan dengan menggabungkan bukti dokumen, hasil wawancara, dan analisis isi instrumen. Literatur metodologi assessment modern merekomendasikan strategi serupa untuk studi alignment dan validasi lokal terhadap kerangka global.

Dengan latar empiris dan landasan teoritis tersebut, penelitian ini bertujuan untuk mendeskripsikan pelaksanaan TOAFL di UIN Syarif Hidayatullah Jakarta serta menganalisis tingkat kesesuaian struktur, indikator, dan kriteria penilaian TOAFL dengan CEFR dan ACTFL. Selain itu, penelitian dimaksudkan menghasilkan rekomendasi praktis bagi pembuat kebijakan akademik dan pengelola pusat bahasa—termasuk rekomendasi teknis untuk penyusunan blueprint tes, rubrik penilaian performa, dan program peningkatan kompetensi penilai—agar TOAFL dapat berfungsi sekaligus sebagai instrumen penilaian pembelajaran yang valid dan alat sertifikasi yang lebih portabel. Temuan yang dihasilkan diharapkan memberikan kontribusi teoretis pada studi evaluasi bahasa Arab dan menyediakan dasar empiris bagi harmonisasi standar penilaian di konteks perguruan tinggi Islam.

Secara ringkas, penelitian ini mengisi celah penting dalam literatur penilaian bahasa di Indonesia dengan menyajikan studi komprehensif tentang hubungan antara alat penilaian lokal (TOAFL) dan kerangka kompetensi internasional (CEFR, ACTFL). Signifikansi penelitian mencakup aspek teoretis—memperkaya kajian alignment dan validasi konstruk dalam assessment bahasa Arab—serta aspek praktis—memberi pedoman perbaikan desain tes dan pelatihan penilai demi peningkatan mutu pendidikan bahasa Arab di lingkungan kampus. Mengingat pembaruan dan tuntutan maksimalitas kompetensi berbahasa di level internasional, upaya harmonisasi seperti yang diusulkan oleh penelitian ini menjadi langkah penting dan tepat waktu.

KAJIAN TEORI

1. Konsep Evaluasi Pembelajaran Bahasa

Evaluasi pembelajaran bahasa merupakan komponen integral dalam sistem pendidikan yang berfungsi menilai efektivitas proses pembelajaran serta mengukur tingkat pencapaian kompetensi peserta didik. Dalam teori modern, evaluasi tidak lagi dipandang sekadar sebagai alat seleksi, melainkan sebagai instrumen refleksi pedagogis yang membantu perbaikan kurikulum, metode, dan strategi pengajaran., penilaian bahasa yang efektif harus bersifat komunikatif, autentik, dan berorientasi pada performa nyata, bukan semata-mata pada penguasaan bentuk linguistik. Dalam kerangka ini, evaluasi berfungsi sebagai mekanisme umpan balik (*feedback mechanism*) yang mendukung pembelajaran berkelanjutan menekankan pentingnya *construct validity* dan *usefulness*, yakni sejauh mana hasil tes benar-benar mengukur kemampuan berbahasa yang relevan dan dapat digunakan untuk tujuan pendidikan maupun administratif.

Evaluasi pembelajaran bahasa dapat diklasifikasikan menjadi tiga bentuk utama: formatif, sumatif, dan diagnostik. Evaluasi formatif dilakukan selama proses pembelajaran untuk mengidentifikasi kekuatan dan kelemahan siswa; sumatif dilakukan pada akhir periode belajar untuk menentukan pencapaian kompetensi; sedangkan diagnostik digunakan untuk memetakan profil kemampuan sebelum pembelajaran dimulai. Ketiga bentuk evaluasi tersebut menuntut alat ukur yang valid, reliabel, dan adil (*fair*). Dalam konteks penilaian bahasa Arab, prinsip keadilan dan representativitas materi menjadi sangat penting karena bahasa ini memiliki variasi dialek dan konteks penggunaan yang beragam. Hal tersebut menegaskan perlunya sistem tes yang tidak hanya menilai aspek struktur, tetapi juga kompetensi pragmatik dan sosiolinguistik.

2. Prinsip Validitas, Reliabilitas, dan Keadilan dalam Tes Bahasa

Dalam kajian pengukuran bahasa modern, validitas dianggap sebagai inti dari teori evaluasi. Kane menjelaskan bahwa validitas bukan sekadar karakteristik tes, melainkan argumentasi berbasis bukti tentang interpretasi skor yang dihasilkan. Dengan demikian, analisis kesesuaian TOAFL terhadap CEFR dan ACTFL juga merupakan upaya validasi konstruk secara eksternal (*external validation*), karena melibatkan perbandingan konsep kompetensi yang diukur dengan standar global. Sementara itu, reliabilitas mengacu pada konsistensi hasil tes, sangat dipengaruhi oleh faktor manusia (*rater*) dan kualitas butir soal. Tes yang reliabel menuntut prosedur penilaian yang transparan, rubrik yang jelas, dan pelatihan penguji yang memadai.

Keadilan (*fairness*) menjadi isu etis yang semakin menonjol dalam teori evaluasi bahasa terkini., keadilan dalam asesmen mencakup aksesibilitas materi tes, kesetaraan konteks, serta perlakuan yang tidak bias terhadap latar sosial dan linguistik peserta. Dalam konteks TOAFL, isu keadilan dapat muncul apabila konten tes lebih menonjolkan aspek gramatikal yang tidak proporsional dibandingkan kemampuan komunikatif, atau jika desain instrumen kurang representatif terhadap variasi kemampuan mahasiswa non-latar Arab. Oleh karena itu, evaluasi kesesuaian TOAFL dengan standar CEFR dan ACTFL sekaligus berfungsi untuk memastikan bahwa sistem tes lokal memenuhi prinsip validitas, reliabilitas, dan *fairness* secara teoritis dan praktis.

3. Teori Modern Pengukuran Kemampuan Berbahasa

Pendekatan modern terhadap pengukuran bahasa berakar pada teori *Communicative Language Testing* (CLT) yang berkembang sejak 1980-an dan mengalami reformulasi signifikan dalam dua dekade terakhir. CLT menekankan bahwa kompetensi berbahasa tidak dapat diukur hanya melalui pemahaman struktur atau kosakata, melainkan melalui kemampuan menggunakan bahasa dalam konteks komunikasi nyata. Menurut Weir dan O’Sullivan (2022), pengukuran bahasa modern harus berlandaskan *socio-cognitive framework*, yang mengintegrasikan faktor konteks, kognisi, dan interaksi sosial dalam penyusunan instrumen tes. Model ini diadopsi secara luas oleh lembaga-lembaga pengembang tes internasional seperti Cambridge Assessment dan ACTFL untuk memastikan bahwa setiap butir tes memiliki *contextual relevance* dan *cognitive authenticity*.

Dalam paradigma ini, tes bahasa harus menilai empat keterampilan utama—menyimak, berbicara, membaca, dan menulis—dalam konteks yang saling berkaitan. Bachman (2019) menyebut hal ini sebagai *integrated performance assessment*, yaitu pengujian yang mensimulasikan aktivitas komunikasi nyata dan menilai kemampuan peserta dalam menggunakan bahasa untuk mencapai tujuan tertentu. Penilaian semacam ini berbeda secara prinsipil dengan tes berbasis pilihan ganda yang hanya mengukur pengetahuan linguistik terpisah dari penggunaan aktual. Dengan demikian, pendekatan teori pengukuran bahasa modern memberikan kerangka konseptual yang tepat untuk meninjau apakah TOAFL di UIN Jakarta telah bergerak menuju evaluasi kompetensi komunikatif sebagaimana ditekankan dalam CEFR dan ACTFL.

4. Kerangka Standar Bahasa Internasional (CEFR)

Common European Framework of Reference for Languages (CEFR) merupakan sistem klasifikasi kompetensi bahasa yang dikembangkan oleh *Council of Europe* dan telah menjadi acuan internasional sejak tahun 2001. Versi terbaru, *CEFR Companion Volume* (2020), memperluas deskriptor kompetensi dengan menambahkan dimensi mediasi, interaksi daring, dan plurilingualism. CEFR mengelompokkan kemampuan bahasa ke dalam enam level—A1 hingga C2—yang menggambarkan perkembangan kemampuan pengguna bahasa dari tahap pemula hingga mahir. Setiap level dirinci dalam deskriptor performatif yang mengukur kemampuan melakukan tindakan komunikatif dalam konteks sosial dan profesional.

Penerapan CEFR tidak hanya sebagai alat penilaian, tetapi juga sebagai dasar penyusunan kurikulum dan instrumen pembelajaran. menegaskan bahwa keselarasan antara tes dan CEFR (*test alignment*) merupakan indikator penting validitas konstruk dan konsistensi antar-institusi. Oleh karena itu, ketika sistem tes bahasa Arab seperti TOAFL hendak disetarakan dengan CEFR, diperlukan analisis mendalam terhadap *can-do statements*, jenis teks yang diujikan, serta kompleksitas tugas komunikatif. Keselarasan ini penting agar hasil TOAFL dapat dipetakan ke level CEFR yang diakui secara internasional, sehingga meningkatkan daya saing dan portabilitas hasil tes bagi mahasiswa UIN Jakarta di ranah akademik global.

5. Kerangka Standar Bahasa Internasional (ACTFL)

Selain CEFR, sistem penilaian lain yang berpengaruh luas adalah pedoman dari *American Council on the Teaching of Foreign Languages* (ACTFL). ACTFL Proficiency Guidelines versi revisi tahun 2024 menegaskan pentingnya *performance-based assessment* dengan klasifikasi kemampuan ke dalam lima tingkat utama: Novice, Intermediate, Advanced, Superior, dan Distinguished. Setiap tingkat dibagi lagi menjadi sublevel Low, Mid, dan High yang

menggambarkan variasi performa dalam setiap domain keterampilan. ACTFL menilai kemampuan berdasarkan kriteria FACT—Function, Accuracy, Content, dan Text type—yang digunakan untuk mengukur sejauh mana peserta dapat berinteraksi secara spontan, akurat, dan bermakna.

ACTFL berperan penting dalam pendidikan bahasa karena memberikan acuan bagi penyusunan rubrik penilaian dan pelatihan penguji. Penelitian terbaru menunjukkan bahwa penerapan ACTFL meningkatkan konsistensi penilaian antar-rater dan memberikan dasar empiris bagi pengembangan tes berbasis performa. Dalam konteks evaluasi TOAFL, ACTFL dapat menjadi kerangka pembandingan yang tepat karena sama-sama menilai keterampilan reseptif dan produktif. Analisis kesesuaian TOAFL terhadap ACTFL akan membantu mengidentifikasi apakah rubrik dan indikator penilaian di TOAFL telah merepresentasikan performa komunikatif yang sejalan dengan praktik penilaian bahasa modern.

6. TOAFL (Test of Arabic as a Foreign Language): Konsep dan Praktik

TOAFL dikembangkan oleh Pusat Bahasa UIN Syarif Hidayatullah Jakarta sebagai instrumen standar untuk mengukur kemampuan bahasa Arab mahasiswa dan masyarakat umum. Tes ini terdiri atas beberapa komponen utama, antara lain *listening comprehension*, *structure and written expression*, *reading comprehension*, serta kadang disertai *writing* atau *speaking test*. Berdasarkan dokumen resmi Pusat Bahasa UIN (2021), TOAFL dirancang untuk menilai kemampuan bahasa Arab akademik dan umum, namun orientasi utama masih dominan pada kemampuan struktural dan pemahaman teks tertulis.

Beberapa penelitian lokal, seperti Hidayat (2022) dan Nuryani (2023), menunjukkan bahwa TOAFL memiliki reliabilitas internal yang baik, namun validitas konstruksinya masih perlu diperkuat karena keterbatasan dalam menilai performa komunikatif. Struktur soal yang didominasi pilihan ganda menyebabkan dimensi pragmatik dan interaksi sosial kurang terwakili. Dalam konteks ini, penelitian kualitatif diperlukan untuk menilai apakah indikator dan konten TOAFL telah selaras dengan *communicative descriptors* dalam CEFR dan ACTFL, serta untuk mengidentifikasi area yang memerlukan perbaikan.

7. Studi Terdahulu dan Kesenjangan Penelitian

Kajian tentang evaluasi bahasa Arab di Indonesia umumnya berfokus pada efektivitas pembelajaran atau reliabilitas tes, bukan pada kesesuaian dengan standar global. Misalnya, studi oleh Ahmad & Mulyadi (2022) membahas strategi evaluasi pembelajaran bahasa Arab di perguruan tinggi, tetapi tidak menyinggung kesetaraan standar internasional. Sebaliknya, penelitian oleh O'Sullivan (2022) dan Fulcher (2023) di konteks Eropa dan Amerika menunjukkan pentingnya proses *alignment* agar hasil tes lokal dapat dibandingkan dengan CEFR atau ACTFL. Perbandingan ini menunjukkan adanya *research gap* dalam konteks Indonesia, khususnya terkait tes kemampuan bahasa Arab.

Kesenjangan penelitian ini penting karena tanpa adanya pemetaan yang jelas, hasil TOAFL tidak dapat diinterpretasikan secara universal dan berisiko kehilangan makna dalam konteks internasionalisasi pendidikan tinggi Islam. Dengan demikian, penelitian ini bukan sekadar studi teknis, tetapi juga kontributif terhadap kebijakan akademik nasional yang menuntut standar mutu dan portabilitas hasil asesmen bahasa. Penelitian ini menutup celah tersebut dengan menganalisis TOAFL secara komparatif terhadap CEFR dan ACTFL melalui pendekatan kualitatif deskriptif.

8. Relevansi Kajian terhadap Pengembangan Pembelajaran Bahasa Arab

Penilaian yang selaras dengan standar global memiliki dampak luas terhadap kurikulum dan praktik pembelajaran. Hasil penelitian alignment dapat digunakan untuk meninjau ulang kurikulum berbasis kompetensi, mendesain bahan ajar yang menargetkan level CEFR tertentu, dan melatih dosen agar mengintegrasikan deskriptor ACTFL ke dalam aktivitas kelas. Dalam konteks UIN Jakarta, evaluasi kesesuaian TOAFL diharapkan mendorong pergeseran paradigma dari tes berbasis struktur menuju *competency-based testing* yang lebih autentik dan komunikatif.

Lebih jauh, hasil penelitian ini berpotensi menjadi model bagi perguruan tinggi Islam lain dalam memperbarui sistem evaluasi bahasa Arabnya. Dengan mengadopsi prinsip CEFR dan ACTFL, institusi pendidikan tinggi dapat meningkatkan daya saing lulusan, memperluas pengakuan hasil tes secara global, dan memperkuat posisi bahasa Arab sebagai bahasa akademik modern yang sejajar dengan bahasa internasional lain.

METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif deskriptif dalam kerangka studi kasus karena fokusnya adalah memahami secara mendalam pelaksanaan instrumen tes (TOAFL) di UIN Syarif Hidayatullah Jakarta serta kesesuaiannya dengan standar internasional (CEFR dan ACTFL). Pendekatan kualitatif dipilih karena memungkinkan penelusuran konteks, makna dan persepsi pengelola serta peserta tes yang tidak dapat ditangkap melalui metode kuantitatif semata. Data dikumpulkan dan dianalisis secara sistematis untuk membangun pemahaman holistik tentang fenomena yang diteliti.

Sebagai penelitian kualitatif studi kasus, hasil temuan bersifat kontekstual dan tidak bermaksud digeneralisasi ke seluruh perguruan tinggi Islam di Indonesia. Namun, melalui penjelasan konteks yang kaya, pembaca dapat melakukan transferabilitas ke situasi serupa.

HASIL DAN PEMBAHASAN

1. Deskripsi Sistem Tes TOAFL di UIN Syarif Hidayatullah Jakarta

Hasil pengumpulan data menunjukkan bahwa TOAFL di UIN Syarif Hidayatullah Jakarta diselenggarakan oleh Pusat Bahasa sebagai salah satu bentuk sertifikasi kemampuan bahasa Arab bagi mahasiswa dan dosen. Tes ini umumnya terdiri atas tiga hingga empat komponen utama, yaitu *listening comprehension*, *structure and written expression*, *reading comprehension*, serta pada beberapa edisi ditambahkan *writing test*. Tes tersebut dilaksanakan dalam konteks administratif dan akademik—yakni sebagai syarat kelulusan, beasiswa, dan seleksi program pascasarjana. Struktur ini mengadaptasi model TOEFL sebagai referensi awal, sehingga pendekatannya lebih bersifat *discrete-point* dan *grammar-based*, bukan berbasis kompetensi komunikatif secara utuh.

Dari wawancara dengan pengelola tes, diketahui bahwa proses penyusunan soal dilakukan oleh tim dosen bahasa Arab internal, tanpa menggunakan kerangka standar internasional eksplisit seperti CEFR atau ACTFL. Meskipun terdapat upaya menjaga konsistensi tingkat kesulitan dan reliabilitas antarversi, mekanisme *test equating* belum dilakukan secara sistematis. Hal ini menyebabkan variasi pada distribusi skor antarperiode. Pengelola menyatakan bahwa instrumen TOAFL masih berorientasi pada penguasaan

struktur gramatikal dan pemahaman teks tertulis, sedangkan aspek lisan (komunikatif dan performatif) belum terakomodasi dengan optimal.

Selain itu, sistem skoring TOAFL masih berbasis jumlah jawaban benar (raw score) yang dikonversi ke rentang 310–677, mengikuti model TOEFL Paper-Based Test. Tidak terdapat deskripsi performatif yang menjelaskan kemampuan peserta berdasarkan level tertentu, seperti yang digunakan dalam CEFR (A1–C2) maupun ACTFL (Novice–Superior). Akibatnya, interpretasi hasil tes bersifat normatif (siapa lebih tinggi), bukan kriteria (apa yang bisa dilakukan penutur pada level tersebut). Kondisi ini mengindikasikan adanya *mismatch* antara paradigma pengujian lokal dan paradigma kompetensi global yang berorientasi pada kemampuan berkomunikasi nyata (*communicative competence*).

2. Analisis Kesesuaian TOAFL dengan CEFR dan ACTFL

Analisis isi instrumen TOAFL terhadap deskriptor CEFR dan ACTFL menunjukkan bahwa sebagian besar butir soal TOAFL hanya mencakup aspek *linguistic competence* — meliputi morfologi, sintaksis, dan pemahaman teks literal. Sebaliknya, CEFR dan ACTFL menekankan integrasi antara *linguistic*, *sociolinguistic*, dan *pragmatic competence* (Council of Europe, 2020). Dengan demikian, TOAFL baru mewakili sekitar dua dari enam skala kemampuan CEFR, yakni A1–B1, terutama dalam keterampilan membaca dan mendengarkan literal, sementara kemampuan berinteraksi, menulis ekspresif, dan berbicara spontan belum terukur.

Dalam perspektif ACTFL, TOAFL cenderung berkorelasi dengan rentang level *Novice High* hingga *Intermediate Mid*, di mana peserta dapat memahami wacana sederhana yang terstruktur, tetapi belum mencapai kemampuan menyusun teks fungsional atau bernegosiasi makna dalam konteks sosial. Analisis komparatif juga memperlihatkan bahwa TOAFL belum menggunakan rubrik deskriptif performatif sebagaimana disarankan ACTFL Performance Descriptors, yang menilai kemampuan peserta berdasarkan *task-based outcome*, bukan sekadar skor pilihan ganda. Ketidaksesuaian ini menunjukkan bahwa TOAFL masih berada pada paradigma *language knowledge testing*, bukan *language use assessment*.

Namun demikian, TOAFL memiliki potensi besar untuk dikembangkan menjadi instrumen berstandar internasional. Struktur tes yang relatif stabil, sistem administrasi yang mapan, dan basis data peserta yang luas dapat menjadi fondasi kuat bagi proses *alignment* atau penyetaraan dengan CEFR/ACTFL. Beberapa langkah adaptif dapat dilakukan, seperti pengembangan *can-do descriptors* untuk setiap level skor, uji validasi empiris menggunakan *item response theory (IRT)*, serta pelatihan *rater* untuk tes lisan dan tulisan sesuai standar global. Dengan demikian, TOAFL dapat bertransformasi dari tes akademik internal menjadi instrumen pengukuran kompetensi bahasa Arab yang diakui secara nasional maupun internasional.

3. Implikasi terhadap Pembelajaran Bahasa Arab di Perguruan Tinggi

Temuan di atas memiliki implikasi langsung terhadap pembelajaran bahasa Arab di perguruan tinggi Islam, terutama dalam hal desain kurikulum, strategi pembelajaran, dan asesmen berbasis kompetensi. Ketika sistem evaluasi masih menitikberatkan pada kemampuan struktural, maka proses pembelajaran pun cenderung mengarah pada hafalan kaidah dan latihan bentuk-bentuk gramatikal. Hal ini berpotensi menurunkan kemampuan komunikatif mahasiswa, yang justru menjadi fokus utama CEFR dan ACTFL. Oleh karena itu, pembelajaran bahasa Arab di UIN dan perguruan tinggi Islam lainnya perlu bergeser menuju pendekatan *proficiency-oriented instruction*, yang menekankan pada kemampuan menggunakan bahasa dalam konteks sosial dan akademik nyata.

Pusat Bahasa juga perlu mengintegrasikan hasil penelitian ini ke dalam pelatihan pengajar dan penguji. Pemahaman terhadap deskriptor CEFR/ACTFL dapat membantu dosen menetapkan tujuan pembelajaran yang lebih jelas, seperti “mahasiswa mampu memperkenalkan diri dan menjelaskan kegiatan akademik dalam bahasa Arab pada level A2–B1” daripada sekadar “mahasiswa memahami struktur fi‘il dan isim.” Selain itu, kurikulum perlu menyesuaikan urutan materi dengan perkembangan level kompetensi, sebagaimana dianjurkan oleh Council of Europe dan ACTFL (2022). Dengan demikian, antara pembelajaran, penilaian, dan tes sertifikasi akan terjalin hubungan yang konsisten dan berkelanjutan.

4. Pembahasan Integratif

Secara teoretis, hasil penelitian ini memperkuat pandangan bahwa tes kemampuan bahasa yang efektif harus mengukur lebih dari sekadar pengetahuan linguistik; ia harus menilai kemampuan penggunaan bahasa dalam konteks sosial dan pragmatik. TOAFL UIN Jakarta, meskipun telah menjadi instrumen evaluasi yang mapan, masih perlu menyesuaikan diri dengan prinsip validitas konstruktif (*construct validity*) yang menuntut keterpaduan antara format tes, tujuan pengajaran, dan deskriptor kemampuan. Dalam kerangka teori *Communicative Language Testing* (Bachman & Palmer, 2010), validitas TOAFL dapat meningkat apabila indikator dan tugas uji mencerminkan situasi komunikasi nyata yang relevan dengan konteks akademik peserta.

Temuan ini juga mengonfirmasi adanya kesenjangan antara *assessment for selection* dan *assessment for learning*. Selama ini TOAFL lebih digunakan untuk penentuan kelulusan atau seleksi administratif, bukan sebagai alat diagnostik yang membantu mahasiswa memahami keunggulan dan kelemahan kemampuan berbahasa mereka. Konversi TOAFL menuju sistem berbasis level CEFR/ACTFL akan membuka peluang untuk menjadikannya alat *formative feedback* yang mendorong perbaikan pembelajaran. Oleh karena itu, hasil penelitian ini tidak hanya penting bagi UIN Jakarta, tetapi juga bagi jaringan perguruan tinggi Islam lain yang menggunakan TOAFL sebagai acuan.

KESIMPULAN

Penelitian ini bertujuan untuk menganalisis sistem dan pelaksanaan Tes Kemampuan Bahasa Arab (TOAFL) di UIN Syarif Hidayatullah Jakarta serta menilai sejauh mana kesesuaiannya dengan standar internasional CEFR (Common European Framework of Reference for Languages) dan ACTFL (American Council on the Teaching of Foreign Languages). Berdasarkan hasil temuan lapangan dan analisis komparatif, dapat disimpulkan bahwa TOAFL memiliki struktur kelembagaan dan sistem administrasi yang mapan, namun masih berorientasi pada paradigma *language knowledge testing*, bukan *language use assessment*. Tes ini menitikberatkan pada aspek struktural dan pemahaman teks tertulis, sementara kemampuan komunikatif dan performatif belum terukur secara memadai.

Analisis terhadap format dan konten TOAFL menunjukkan bahwa sebagian besar butir soal hanya mencakup dimensi *linguistic competence*, sementara dimensi *sociolinguistic* dan *pragmatic competence* yang menjadi inti dari CEFR dan ACTFL belum tercakup secara sistematis. Akibatnya, hasil TOAFL belum dapat dikonversikan secara langsung ke level kompetensi CEFR (A1–C2) atau kategori ACTFL (Novice–Superior). Meskipun demikian, struktur dan mekanisme tes yang sudah stabil menjadi potensi awal bagi pengembangan instrumen yang lebih komprehensif, terutama jika dilakukan *alignment* dengan deskriptor global.

Temuan ini mengindikasikan adanya kesenjangan paradigma antara tes kemampuan bahasa Arab yang bersifat lokal dan kerangka pengukuran internasional yang berbasis performatif. TOAFL masih menilai apa yang diketahui peserta tentang bahasa, bukan apa yang dapat dilakukan dengan bahasa tersebut. Padahal, teori *Communicative Language Testing* (Bachman & Palmer, 2010; Fulcher, 2021) menegaskan bahwa validitas sebuah tes bahasa bergantung pada sejauh mana ia mengukur kemampuan komunikasi autentik yang relevan dengan konteks sosial pengguna bahasa. Dalam konteks pendidikan tinggi Islam, kesenjangan ini berimplikasi langsung terhadap mutu pembelajaran bahasa Arab yang masih dominan berbasis struktur, bukan kompetensi komunikatif.

Secara teoretis, penelitian ini memperkuat relevansi penerapan *proficiency-based assessment* dalam pengajaran bahasa Arab di Indonesia. Standar CEFR dan ACTFL terbukti memberikan deskriptor yang jelas, terukur, dan dapat diterapkan lintas konteks untuk menggambarkan perkembangan kemampuan bahasa secara berjenjang. Dengan demikian, hasil penelitian ini menegaskan pentingnya mengintegrasikan standar tersebut dalam sistem evaluasi lokal agar penilaian kemampuan bahasa Arab di perguruan tinggi tidak hanya memiliki nilai administratif, tetapi juga fungsional dalam konteks akademik dan profesional.

REKOMENDASI

Berdasarkan temuan dan kesimpulan di atas, penelitian ini merekomendasikan tiga arah pengembangan utama: (a) penguatan instrumen tes, (b) integrasi standar internasional, dan (c) reformulasi kebijakan pembelajaran bahasa Arab.

Pertama, penguatan instrumen TOAFL perlu dilakukan melalui *alignment* struktural dan konseptual terhadap CEFR dan ACTFL. Hal ini dapat dimulai dengan menyusun peta indikator kemampuan (*can-do descriptors*) untuk setiap rentang skor TOAFL, sehingga hasil tes dapat diinterpretasikan secara deskriptif, bukan hanya numerik. Selain itu, perlu dikembangkan komponen tes lisan (*speaking*) dan tulisan (*writing*) dengan rubrik penilaian performatif yang melibatkan pelatihan penguji bersertifikat. Penggunaan model *Item Response Theory* (IRT) juga disarankan untuk menjamin validitas empiris dan kesetaraan tingkat kesulitan antarversi tes.

Kedua, integrasi standar internasional perlu diwujudkan melalui kolaborasi antara lembaga pengelola TOAFL dan institusi yang memiliki pengalaman dalam *test standardization*, seperti Cambridge Assessment atau ACTFL Council. Kolaborasi ini dapat mencakup lokakarya penyetaraan (*benchmarking*), penyusunan panduan uji kompetensi berbasis tugas (*task-based testing*), serta pelatihan pengembang soal sesuai dengan prinsip *communicative language testing*. Dengan demikian, TOAFL dapat bertransformasi menjadi instrumen nasional yang memiliki validitas global tanpa kehilangan konteks keislaman dan keindonesiaannya.

Ketiga, dalam ranah pembelajaran, hasil penelitian ini menegaskan pentingnya *assessment literacy* di kalangan pengajar bahasa Arab. Dosen dan guru perlu memahami hubungan antara deskriptor kompetensi, tujuan pembelajaran, dan jenis penilaian. Penerapan pendekatan *proficiency-oriented instruction* akan membantu menciptakan kesinambungan antara pengajaran, evaluasi, dan sertifikasi kemampuan bahasa. Pergeseran ini tidak hanya akan meningkatkan kualitas lulusan dalam aspek komunikatif, tetapi juga memperkuat posisi bahasa Arab sebagai bahasa akademik dan global.

Selain itu, penelitian ini membuka ruang bagi studi lanjutan yang lebih mendalam, terutama dalam menguji validitas empiris TOAFL dengan metode kuantitatif, seperti *factor analysis* atau *Rasch modeling*, serta studi perbandingan antarperguruan tinggi yang telah melakukan adaptasi CEFR/ACTFL. Kajian

longitudinal juga diperlukan untuk menilai dampak perubahan sistem tes terhadap perkembangan kemampuan mahasiswa dalam jangka panjang. Dengan demikian, hasil penelitian ini diharapkan menjadi langkah awal menuju standarisasi nasional pengukuran kemampuan bahasa Arab di Indonesia yang berorientasi pada kualitas dan daya saing global.

DAFTAR PUSTAKA

- ACTFL. (2022). *Performance Descriptors for Language Learners*. American Council on the Teaching of Foreign Languages.
- ACTFL. (2024). *ACTFL Proficiency Guidelines (2024 revision)*. <https://www.actfl.org/news/revised-actfl-proficiency-guidelines-released>. ACTFL+1
- ACTFL. (2024). *ACTFL Proficiency Guidelines (Revised Edition)*. American Council on the Teaching of Foreign Languages. <https://www.actfl.org>
- American Council on the Teaching of Foreign Languages (ACTFL). (2024). *ACTFL Proficiency Guidelines (2024 revision)*. <https://www.actfl.org/news/revised-actfl-proficiency-guidelines-released>. ACTFL+1
- Bachman, L. F., & Palmer, A. S. (2010). *Language Assessment in Practice: Developing Language Assessments and Justifying Their Use in the Real World*. Oxford University Press.
- Brown, H. D., & Abeywickrama, P. (2021). *Language Assessment: Principles and Classroom Practices* (3rd ed.). Pearson.
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment — Companion Volume* (definitive version 2020; pembahasan lanjutan dan aplikasi tersedia 2021–2022). <https://www.coe.int/en/web/common-european-framework-reference-languages>. Portal+1
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Companion Volume*. Strasbourg: Council of Europe Publishing.
- Fulcher, G. (2021). *Practical Language Testing* (3rd ed.). Routledge.
- Fulcher, G., & Harding, L. (Eds.). (2021). *The Routledge Handbook of Language Testing* (2nd ed.). Routledge. <https://www.routledge.com/The-Routledge-Handbook-of-Language-Testing-2nd-Edition/Fulcher-Harding/p/book/9780367428818>. Taylor & Francis
- Fulcher, G., & Harding, L. (2021). *The Routledge Handbook of Language Testing* (2nd ed.). Routledge.
- Hidayat, A. (2019). “Analisis Validitas dan Reliabilitas Tes TOAFL di Perguruan Tinggi Keagamaan Islam.” *Jurnal Pendidikan Bahasa Arab*, 6(2), 145–160.
- Hidayat, A. (2022). “Analisis Validitas Tes TOAFL di Perguruan Tinggi Islam.” *Al-Arabiya: Jurnal Bahasa Arab*, 8(2), 134–148.
- Kane, M. (2022). *Validation in Educational Measurement: An Argument-Based Approach*. Routledge.
- Liskin-Gasparro, J. (2023). “Performance Assessment in Language Testing: Revisiting the ACTFL Paradigm.” *Language Testing Review*, 41(1), 45–61.
- Mukhtar, A. (2021). *Evaluasi Pembelajaran Bahasa Arab di Perguruan Tinggi Islam: Teori dan Aplikasi*. Jakarta: Kencana.
- O’Sullivan, B. (2022). *The CEFR and Language Testing Alignment*. Cambridge University Press.
- O’Sullivan, B., Figueras, N., & Little, D. (2022). *Aligning Language Education with the CEFR: Handbook and Commentary*. Cambridge University Press.
- Pusat Pengembangan Bahasa UIN Syarif Hidayatullah Jakarta. (2020–2021). *Pengumuman dan dokumentasi teknis pelaksanaan TOAFL/ETIC/TIPA (halaman administratif dan pedoman pendaftaran tes)*. <https://pusatbahasa.uinjkt.ac.id>. pusatbahasa.uinjkt.ac.id+1
- Pusat Pengembangan Bahasa UIN Syarif Hidayatullah Jakarta. (2021). *Pedoman Pelaksanaan Tes TOAFL*. UIN Jakarta.
- Saville, N., & Green, A. (2024). *Language Testing and Policy: Current Trends and Future Directions*. Cambridge University Press.