

Problematika Evaluasi dalam Pembelajaran BIPA Tingkat Madya: Studi Kualitatif di Universitas Walailak, Thailand

Mahardhika Agung Wicaksono¹, Ahmad Jami'ul Amil², Siti Mutiatun³

¹²³ Universitas Trunojoyo Madura; Indonesia

Correspondence e-mail*, agungmahar87@gmail.com¹, ahmadamil@trunojoyo.ac.id², siti.mutiatun@trunojoyo.ac.id³

Submitted:2025/12/10 Revised: 2025/12/20 Accepted: 2025/12/24 Published: 2025/12/26

Abstract Evaluation plays a crucial role in Indonesian as a Foreign Language (BIPA) learning, particularly at the intermediate level, which represents a transitional stage from basic language acquisition toward pragmatic, cultural, and academic language competence. This qualitative study aims to identify and analyze evaluation issues based on teaching materials and learning materials for intermediate-level BIPA at Walailak University, Thailand. The research method used a case study approach with data collection techniques through participatory observation, in-depth interviews with three BIPA teachers, document analysis (syllabus, lesson plans, evaluation sheets, and teaching materials), and focus group discussions with 15 intermediate level students. The findings reveal four main problems: (1) the gap between teaching materials and evaluation instruments; (2) the dominance of product-based summative evaluation over process-based evaluation; (3) challenges in evaluating pragmatic and cultural competencies; and (4) limitations in contextualizing evaluation materials to the Thai reality. These problems stem from the limited availability of teaching materials, the administrative burden on teachers, and the need for more specific assessment training for the Thai context. This study recommends the development of a continuous assessment model integrated with teaching materials based on the local Thai context, as well as assessment literacy training for BIPA teachers.

Keywords Evaluation, Intermediate Level BIPA, Teaching Materials, Walailak University,



© 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY SA) license, <https://creativecommons.org/licenses/by-sa/4.0/>.

PENDAHULUAN

Terbatasnya kajian yang secara spesifik mengkaji problematika evaluasi pembelajaran BIPA tingkat madya dalam internasional khususnya di perguruan tinggi luar negeri seperti Thailand, yang memiliki latar sosial-budaya dan tujuan pembelajaran berbeda dari Indonesia. Sebagian besar penelitian BIPA sebelumnya lebih banyak berfokus pada pengembangan bahan ajar atau evaluasi di tingkat dasar serta dilakukan dalam konteks domestik, sehingga belum menggambarkan secara mendalam kesesuaian instrumen evaluasi dengan kompetensi pragmatik, kultural, dan akademik yang dituntut pada tingkat madya. Maka, masalah utama yang dikaji dalam penelitian ini mengenai ketidaksesuaian antara instrumen evaluasi BIPA tingkat madya dengan konteks, kebutuhan, dan capaian kompetensi mahasiswa Thailand, yang berpotensi menyebabkan evaluasi tidak mampu

mengukur kemampuan komunikatif secara utuh dan bermakna.¹

Peminat BIPA di Thailand bukanlah jumlah kecil, melainkan tersebar di banyak universitas, sehingga program ini dapat berkontribusi pada penyebaran bahasa Indonesia dan pemahaman budaya Nusantara di luar negeri. Kondisi ini menjadikan BIPA bukan sekedar kursus bahasa, tetapi entitas penting dalam konteks hubungan antarbangsa dan integrasi regional.² Untuk menggambarkan skala penyebaran BIPA di Thailand, berikut data statistik jumlah mahasiswa yang tercatat sebagai pemelajar bahasa Indonesia di berbagai universitas di Thailand:

Tabel 1. Sebaran Peserta Program BIPA di Perguruan Tinggi Thailand

Universitas (Thailand)	Perkiraan Jumlah Mahasiswa BIPA/Peserta belajar BI	Keterangan/Catatan
Walailak University	10 orang	Data survei BIPA di Thailand 2023
Thammasat University (Bangkok)	104 orang	Institusi yang tercatat sebagai penyelenggara BIPA
Khon Kaen University	193 orang	Universitas di Thailand yang menawarkan kelas BIPA
Prince of Songkla University	85 orang	Termasuk universitas yang memiliki program BIPA

Sumber: Haman Supriyadi et al. (2023).³

Keberadaan data di atas menegaskan bahwa program BIPA di Thailand memiliki basis peserta yang cukup beragam dan tersebar, sehingga memungkinkan latar kontekstual berbeda dari yang ada di Indonesia. Hal ini sekaligus menunjukkan bahwa bagi institusi seperti Walailak University, BIPA bukan sekedar aksesoris akademik, melainkan bagian dari strategi pendidikan dan diplomasi budaya. Tantangan dalam menyusun bahan ajar maupun instrumen evaluasi tentu tidak bisa dianggap enteng perlu penyesuaian spesifik terhadap konteks lokal, latar belakang peserta, dan tujuan pembelajaran. Penelitian tentang evaluasi BIPA, khususnya pada tingkat madya, menjadi

¹ Wulandari, Annisa Eka, and Neneng Konety. "Diplomasi Budaya Indonesia terhadap Vietnam melalui Program Bahasa Indonesia bagi Penutur Asing (BIPA)." *Padjadjaran Journal of International Relations* 6.2 (2024): 241-257.

² Tiawati, Refa Lina. "Bahasa Indonesia di Thailand menjadi media diplomasi kebahasaan dan budaya di ASEAN melalui pengajaran BIPA." *Jurnal Gramatika: Jurnal Penelitian Pendidikan Bahasa dan Sastra Indonesia* 1.1 (2015): 29-44.

³ Haman Supriyadi et al. *"Harmoni Kata: Panduan Belajar Kosakata Indonesia-Thai"*, Jakarta: SEAMEO QITEP in Language, 2023.
https://www.qiteplanguage.org/assets/files/dokumen/Harmoni%20Kata_Panduan%20Belajar%20Kosa%20Kata%20Bahasa%20Indonesia%20Thai-Copy.pdf

sangat relevan.

Tingkat madya dalam BIPA menempati posisi strategis karena di level ini peserta diharapkan melampaui sekadar “kemampuan dasar untuk bertahan hidup (*survival proficiency*)” dan mulai mengembangkan kompetensi yang diperlukan untuk keperluan akademik atau profesional. Bahan ajar di tingkat ini umumnya memperkenalkan teks yang lebih panjang, ragam ungkapan idiomatis, serta unsur budaya dan pragmatik yang lebih kompleks agar pemelajar tidak cuma paham kata per kata, melainkan memahami makna dalam konteks sosial budaya Indonesia.⁴ Sebagian besar bahan ajar BIPA yang beredar dikembangkan dengan asumsi konteks kehidupan di Indonesia: latar budaya, keseharian, dan norma pragmatik bahasa Indonesia. Dalam pembelajaran di Thailand (dan di kampus seperti Walailak), asumsi tersebut tidak selalu sesuai dengan pengalaman hidup mereka, sehingga mendatangkan tantangan dalam relevansi materi.

Persoalan ini menjadi penting ketika kita mempertimbangkan fungsi evaluasi dalam pembelajaran. Evaluasi tidak hanya berfungsi sebagai alat ukur hasil belajar (*assessment of learning*), melainkan idealnya menjadi bagian dari proses pembelajaran itu sendiri memberi umpan balik, memetakan kebutuhan, serta mendorong perbaikan berkelanjutan (*assessment for learning*). Ketika instrumen evaluasi dibangun berdasarkan bahan ajar yang tidak sepenuhnya relevan dengan konteks peserta, maka kemampuan nyata peserta dalam menggunakan bahasa Indonesia secara kontekstual mungkin tidak terukur secara akurat. Evaluasi bisa gagal menangkap kompetensi pragmatik, kultural, dan komunikatif yang sesungguhnya.⁵

Kondisi program BIPA di kampus seperti Walailak University menjadi kasus khas, di satu sisi ada upaya serius untuk memperkenalkan bahasa dan budaya Indonesia, namun di sisi lain bahan ajar dan evaluasi seringkali belum disesuaikan secara memadai dengan latar belakang dan kebutuhan mahasiswa Thailand. Penelitian yang dikembangkan untuk mahasiswa Thailand sering terbatas pada tingkat dasar (A1/A2) dengan fokus kosakata dasar dan situasi sehari-hari, tanpa menyiapkan materi tingkat madya atau lanjutan yang kaya akan konteks budaya maupun akademik.⁶ Kondisi ini memunculkan pertanyaan kritis, apakah bahan ajar dan evaluasi yang tersedia sudah memadai untuk membantu mahasiswa Thailand mencapai kompetensi bahasa

⁴ Septyani, Dwi Endah, Zainal Rafli, and Liliana Muliastuti. "Keterbacaan Wacana Buku Teks BIPA" Sahabatku Indonesia" Tingkat Madya." *Indonesian Language Education and Literature* 6.1 (2020): 13-23.

⁵ Sudarman, Silas, et al. *Evaluasi Pendidikan Abad 21: Inovasi, Spiritualitas, dan Relevansi*. Jambi: PT. Sonpedia Publishing Indonesia, 2025.

⁶ Wibawani, Sri. *Pengembangan Instrumen Tes Berbasis Digital untuk Mengukur Kemampuan Membaca Mahasiswa BIPA Tingkat C1*. *Skripsi*. Universitas Brawijaya, 2021.

Indonesia yang mendalam dan relevan, ataukah hanya sekadar memenuhi standar minimal?.

Dalam pedagogis hal tersebut menunjukkan perlunya kajian kritis terhadap “kesesuaian kontekstual (*contextual fit*)” antara materi ajar, latar sosial-budaya peserta, dan instrumen evaluasi.⁷ Kesesuaian kontekstual ini mencakup aspek bahasa (kosakata, struktur, idiom), budaya (referensi budaya, norma komunikasi), serta pragmatik (penggunaan bahasa dalam situasi nyata). Tanpa kesesuaian semacam itu pembelajaran BIPA dapat menjadi mekanik dan terlepas dari realitas penggunaan bahasa sehingga lulusan BIPA mungkin mampu menghadapi tes, tetapi kesulitan menggunakan bahasa dalam interaksi nyata di lingkungan Indonesia atau komunitas berbahasa Indonesia.

Evaluasi BIPA di kampus seperti Walailak bukan hanya soal mengukur kemampuan bahasa, tetapi soal mempersiapkan mahasiswa agar mampu berkomunikasi efektif dalam konteks lintas budaya dan akademik. Jika evaluasi hanya mengukur kemampuan dasar misalnya kosakata dan tata bahasa sederhana maka lulusan BIPA belum tentu siap untuk menulis esai dalam bahasa Indonesia, berdiskusi dengan mahasiswa Indonesia, atau melakukan presentasi akademik dalam bahasa Indonesia. Hal ini akan membatasi potensi BIPA sebagai jembatan bagi kerjasama akademik atau profesional lintas negara.⁸

Penting bagi penelitian tentang evaluasi BIPA tingkat madya di lingkungan internasional untuk mengeksplorasi bagaimana instrumen evaluasi dapat dirancang secara lebih kontekstual dan bermakna misalnya dengan penugasan keterampilan menyimak teks otentik, menulis esai reflektif, praktik wawancara, atau tugas lintas budaya. Pendekatan evaluasi semacam itu tidak hanya mengukur hasil belajar, tetapi juga mendorong pembelajaran mendalam dan relevan dengan kebutuhan nyata mahasiswa.

Penelitian Mintowati dkk. (2021) menyoroti problematika pengajaran BIPA di Universitas Negeri Surabaya dan Walailak University dengan menekankan kendala guru, khususnya keterbatasan bahan ajar dan media pembelajaran, tanpa mengkaji secara spesifik aspek evaluasi pembelajaran.⁹ Sementara penelitian Wigati dan Pramono (2025) melakukan evaluasi program

⁷ Tonji, Nasrulah Ridwan, et al. "Analisis Setting Sosial dalam Penyusunan Bahan Ajar PAI di SMA Negeri 2 Limboto Kabupaten Gorontalo." *Konstruktivisme: Jurnal Pendidikan dan Pembelajaran* 16.2 (2024): 359-370.

⁸ Damawanti, Kustina Dwi Nur, Sri Wahyuni, and Frida Siswiyanti. "Implementasi Multimedia Interaktif Dalam Pembelajaran Menulis Cerita Pada Mahasiswa Asing Di Universitas Islam Malang." *Jurnal Penelitian, Pendidikan, dan Pembelajaran* 20.3 (2025).

⁹ Mintowati, Maria, et al. "Problematics Of BIPA Teachers At Unesa Indonesia And Walailak University Thailand." *International Conference of Humanities and Social Science (Ichss)*. 2021.

pembelajaran BIPA menggunakan model CIPP, namun penelitian tersebut dilakukan pada jenjang pendidikan menengah (SMP) dan bersifat evaluasi program secara umum, bukan pada instrumen evaluasi pembelajaran bahasa di tingkat madya dan konteks perguruan tinggi internasional.¹⁰ Celah penelitian terletak pada belum adanya kajian yang secara khusus dan mendalam membahas problematika evaluasi pembelajaran BIPA tingkat madya di perguruan tinggi luar negeri, terutama terkait kesesuaian instrumen evaluasi dengan kompetensi pragmatik, kultural, dan akademik pemelajar. Penelitian ini menempati posisi strategis untuk mengisi celah tersebut dengan mengkaji secara kualitatif praktik evaluasi BIPA tingkat madya di Universitas Walailak, Thailand, sehingga penting untuk memberikan kontribusi teoretis dan praktis bagi pengembangan evaluasi BIPA yang lebih kontekstual, bermakna, dan relevan dalam konteks internasional.

Kondisi di kampus seperti Walailak University menunjukkan bahwa tantangan evaluasi bukan hanya soal teknis (bagaimana menyusun tes), tetapi soal relevansi, konteks, dan tujuan jangka panjang. Penelitian yang menjadi fokus yaitu “Problematika Evaluasi dalam Pembelajaran BIPA Tingkat Madya” menjadi sangat penting untuk mengungkap sejauh mana evaluasi di BIPA mampu menjembatani antara pembelajaran bahasa sebagai kompetensi linguistik dan sebagai kompetensi sosial-kultural. Hasil penelitian ini tidak hanya relevan bagi Walailak, tetapi juga bagi pengembangan BIPA internasional.

METODE

Penelitian ini menggunakan pendekatan kualitatif dengan desain studi kasus untuk menggali secara mendalam problematika evaluasi dalam pembelajaran BIPA tingkat madya di Universitas Walailak.¹¹ Fokus penelitian diarahkan pada situasi alamiah kelas sehingga temuan yang diperoleh mencerminkan kondisi riil di lapangan. Lokasi penelitian berada di unit penyelenggara program BIPA Universitas Walailak, dengan partisipan tiga pengajar berkewarganegaraan Thailand yang memiliki pengalaman mengajar minimal tiga tahun serta berlatar belakang pendidikan S2 dalam bidang pengajaran bahasa atau linguistik, dan lima belas mahasiswa tingkat madya yang dipilih secara purposif untuk mewakili variasi kemampuan. Strategi pemilihan partisipan ini bertujuan agar sudut pandang pengajar maupun mahasiswa dapat terwakili secara seimbang dalam pemetaan persoalan evaluasi yang terjadi.

¹⁰ Wigati, Prinastining Dyah, and Rudy Pramono. "Evaluasi Program Pembelajaran Bipa Dengan Metode Context, Input, Process, Product (CIPP)." *Media Bina Ilmiah* 19.9 (2025): 5625-5638.

¹¹ Achjar, Komang Ayu Henny, et al. *Metode penelitian kualitatif: Panduan praktis untuk analisis data kualitatif dan studi kasus*. Jambi: PT. Sonpedia Publishing Indonesia, 2023.

Langkah-langkah analisis data dalam penelitian ini dilakukan secara sistematis dengan mengikuti model interaktif Miles, Huberman, dan Saldana. Pertama, seluruh data hasil wawancara, FGD, observasi kelas, dan analisis dokumen ditranskripsi dan dikompilasi, kemudian dilakukan reduksi data dengan cara menyeleksi, memfokuskan, dan mengode informasi yang relevan dengan problematika evaluasi BIPA tingkat madya. Kedua, data yang telah direduksi disajikan dalam bentuk penyajian data (data display) berupa matriks, tabel tematik, dan narasi analitis untuk memudahkan peneliti melihat pola, hubungan, dan kecenderungan antar-temuan dari berbagai sumber. Ketiga, peneliti melakukan penarikan kesimpulan dan verifikasi secara berulang melalui analisis tematik untuk merumuskan makna, kategori, dan tema utama terkait kesesuaian instrumen evaluasi, konteks pembelajaran, serta pengalaman pengajar dan mahasiswa. Keabsahan data dijaga melalui triangulasi sumber dan teknik, yaitu dengan membandingkan temuan dari dokumen, wawancara, FGD, dan observasi, sehingga kesimpulan yang dihasilkan merefleksikan kondisi empiris secara akurat dan dapat dipertanggungjawabkan secara ilmiah.¹²

HASIL DAN PEMBAHASAN

Kesenjangan antara Materi Ajar dan Instrumen Evaluasi

Bahan ajar yang digunakan merupakan kombinasi antara buku ajar BIPA terbitan Indonesia (seperti Sahabatku Indonesia tingkat B1) dan materi tambahan buatan pengajar yang diambil dari sumber online (berita, YouTube). Evaluasi yang dominan terdiri dari Ujian Tengah Semester (UTS) dan Ujian Akhir Semester (UAS) yang bersifat sumatif, serta kuis mingguan. Bentuk soal didominasi pilihan ganda, isian, dan esai pendek untuk keterampilan reseptif (membaca, menyimak), sementara untuk produktif (berbicara, menulis) menggunakan tes performa dengan rubrik.¹³

Dari analisis dokumen dan wawancara menunjukkan adanya *mismatch* antara cakupan materi yang diajarkan dan yang diujikan. Sebagai contoh materi satu bab banyak membahas tentang “Pariwisata Budaya di Indonesia” dengan kosakata spesifik (candi, upacara adat, kain tradisional). Namun dalam soal evaluasi membaca, teks yang digunakan justru tentang “Pariwisata Bahari di Thailand” dengan kosakata yang berbeda meski temanya sama. Hasil jawaban wawancara seorang

¹² Suwarno, Suparjo Adi. "Kinerja kepala madrasah dalam melaksanakan fungsi manajemen pendidikan di MTs Alamiriyah Blokagung Karangdoro Tegalsari Banyuwangi." *Innovative: Journal Of Social Science Research* 3.6 (2023): 7012-7029.

¹³ Pramesti, Riza Dian. "Pengembangan Media Pembelajaran E-Modul (Book Creator) Aktivitas Liburan Menilik Pesona Indonesia Pada Pembelajaran BIPA Tingkat Dasar", *Skripsi*, Universitas Islam Malang, 2022.

pengajar (Responden P2) menyatakan:

“Kadang kami kesulitan membuat soal yang persis sesuai materi karena teks aslinya terlalu panjang atau kompleks. Akhirnya cari teks lain yang lebih sederhana, tapi kosakatanya jadi berbeda. Siswa pun protes, ‘Bu, ini tidak diajarkan’”.

Hal ini menimbulkan ketidakadilan dalam evaluasi dan mengukur kemampuan memahami teks baru, bukan mengukur penguasaan materi yang telah dipelajari. Profil bahan ajar dan evaluasi di program BIPA tingkat madya di Universitas Walailak nyata mencerminkan kombinasi antara buku ajar standar (misalnya buku teks BIPA terbitan Indonesia) dan materi tambahan yang disusun pengajar dari sumber daring seperti artikel, berita, atau video.¹⁴ Evaluasi yang dominan di kelas meliputi UTS, UAS, serta kuis mingguan dengan format soal pilihan ganda, isian, esai pendek untuk keterampilan reseptif, serta tes performa dengan rubrik untuk produktif.

Namun hasil dan wawancara mengungkap adanya ketidaksesuaian antara materi yang diajarkan dan konten evaluasi yang diberikan misalnya materi ajar berfokus pada budaya dan kosakata Indonesia spesifik, sedangkan soal evaluasi memakai teks dengan konteks berbeda. Pengalaman ini sejalan dengan temuan pada pengajaran bahasa asing lain bahwa ketidaksesuaian antara materi dengan asesmen dapat merusak efektivitas pembelajaran. Untuk menegaskan bahwa fenomena *“mismatch”* antara materi dan asesmen bukan hanya masalah lokal tetapi telah diidentifikasi secara akademik, berikut tabel hasil studi literatur yang relevan sebagai data sekunder pendukung,

Tabel 2. Ringkasan Temuan Penelitian Terkait Ketidakesuaian Materi dan Instrumen Evaluasi dalam Pembelajaran Bahasa

Studi/ Penelitian	Fokus Kompetensi / Materi	Bentuk Evaluasi yang Digunakan	Temuan Ketidakesuaian/Dampak terhadap Pembelajaran
Ketidakesuaian Materi dan Asesmen pada Pembelajaran Bahasa Indonesia Teks Negosiasi (Alfiani, Rani & Wahyuni, 2024). ¹⁵	Teks negosiasi memerlukan keterampilan berbicara & interaksi sosial	Tes tertulis / kognitif (soal tertulis)	Asesmen tertulis tidak mengukur kemampuan berbicara/negosiasi secara autentik, sehingga kompetensi produktif siswa rendah

¹⁴ Amil, Ahmad Jami’ul. "Internasionalisasi Bahasa Indonesia Melalui Study Abroad Program Bahasa Indonesia Bagi Penutur Asing (Bipa) Menggunakan Pendekatan Sastra Dan Seni (Studi Kasus Pada Mahasiswa Asean Studies Universitas Walailak Thailand Tahun 2020)." *Prosiding Seminar Internasional Seminar Kepakaran Bahasa Indonesia Untuk Penutur Asing (SEMAR BIPA)*. Vol. 3. No. 1. 2020.

¹⁵ Alfiani, Calvin Reindira, Abdul Rani, and Sri Wahyuni. "Ketidakesuaian Materi Dan Asesmen Pada

Performance-based Assessment as a Current Trend in ELT: Investigating Its Washback Effects on Secondary-School Students' Learning (Sumardi, 2024), ¹⁶	Kemampuan bahasa reseptif dan produktif dalam pembelajaran bahasa asing	Asesmen tradisional (pilihan ganda, esai pendek) vs asesmen performa	Format tradisional terbatas dalam mendeteksi kemampuan faktual & komunikatif; asesmen performa lebih autentik dan memotivasi siswa belajar lebih dalam
Structural relationship between learners' perceptions of a test, learning practices, and learning outcomes: A study on the washback mechanism of a high-stakes test (Dong Manxia, 2020). ¹⁷	Efek tes besar (high-stakes) terhadap praktik belajar dan hasil belajar	Tes besar / high-stakes test	Persepsi siswa terhadap tes mempengaruhi praktik belajar; jika tes tidak valid atau tidak sesuai dengan materi, hasil belajar dan praktik bisa terdistorsi

Tabel ini memperkuat argumen bahwa mismatch materi–evaluasi memengaruhi kualitas pembelajaran dan kompetensi komunikatif siswa, dan bahwa pendekatan tradisional (soal tertulis, pilihan ganda) sering gagal menangkap aspek produktif dan pragmatik bahasa. Dalam program BIPA di Universitas Walailak, hasil wawancara dengan pengajar menunjukkan bahwa kendala terbesar dalam merancang soal yang sesuai materi adalah keterbatasan akses terhadap teks autentik yang sesuai tingkat kesulitan dan konteks budaya, sehingga sering dipilih teks alternatif yang “lebih mudah” namun berbeda dari materi ajar menimbulkan protes dari mahasiswa bahwa soal “tidak diajarkan”. Kondisi ini paralel dengan temuan Alfiani et al. bahwa asesmen formatif atau kognitif (soal tertulis) tidak memadai untuk materi interaksi sosial atau negoisasi.

Pada asesmen berbasis performa dalam pengajaran bahasa (ELT) dapat terlihat bahwa format seperti itu dapat menghasilkan efek positif pada motivasi dan keterlibatan siswa, serta menilai kompetensi lebih autentik dibandingkan tes tradisional. Jika program BIPA di Walailak tetap memakai tes tradisional dominan, maka potensi untuk membentuk kompetensi komunikatif Bahasa Indonesia yang autentik bagi mahasiswa Thailand menjadi terbatas.

Pembelajaran Bahasa Indonesia Teks Negosiasi." *Imajeri: Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 7 (2024).

¹⁶ Sumardi, Sumardi. "Performance-based assessment as a current trend in ELT: Investigating its washback effects on secondary-school students learning." *Kajian Linguistik dan Sastra* 2.1 (2017): 1-11.

¹⁷ Dong, Manxia. "Structural relationship between learners' perceptions of a test, learning practices, and learning outcomes: A study on the washback mechanism of a high-stakes test." *Studies in Educational Evaluation* 64 (2020): 100824.

Pendekatan studi tentang “washback” (efek tes terhadap pembelajaran) juga relevan di sini studi dengan 3.105 siswa di China menunjukkan bahwa persepsi mereka terhadap validitas dan pentingnya tes memengaruhi cara mereka belajar, dan praktik belajar yang dipengaruhi oleh tes berdampak nyata pada hasil belajar.¹⁸ Bila di program BIPA mahasiswa merasakan bahwa evaluasi tidak relevan dengan materi ajar, kemungkinan besar mereka juga akan menyesuaikan cara belajar misalnya belajar “untuk lulus soal” bukan belajar untuk kompetensi sehingga pembelajaran kehilangan arah komunikatif dan budaya.

Kesenjangan antara materi ajar dan evaluasi bukan hanya merupakan fenomena lokal, tetapi bagian dari problem umum dalam assessment bahasa asing jika instrumen tidak disesuaikan secara kontekstual dan komunikatif. Integrasi literatur memperkuat bahwa *mismatch* semacam ini berdampak pada kualitas pembelajaran bahasa, kemampuan produktif, dan relevansi budaya.

Sangat mendesak untuk merancang ulang instrumen evaluasi BIPA: dari dominan tes tertulis dan format tradisional ke asesmen yang lebih autentik misalnya tes performa, tugas komunikasi, asesmen berbasis proyek atau tugas nyata yang relevan dengan konteks mahasiswa agar evaluasi benar-benar mengukur kompetensi bahasa dan budaya sesuai tujuan BIPA tingkat madya dan mendukung kompetensi komunikatif yang nyata bagi mahasiswa internasional.

Dominasi Evaluasi Sumatif Berbasis Produk atas Proses

Sistem evaluasi masih sangat terpusat pada UTS dan UAS (70% nilai akhir). Hasil observasi menunjukkan bahwa aktivitas diskusi, presentasi, dan tugas kecil di kelas meskipun ada, sering kali tidak tercatat secara sistematis sebagai bagian dari penilaian formatif. FGD dengan mahasiswa mengungkapkan kecemasan terhadap ujian akhir yang dianggap “menentukan segalanya”. Mereka merasa bahwa perkembangan mereka selama proses pembelajaran, seperti peningkatan kelancaran berbicara atau partisipasi dalam memahami budaya, tidak terakomodasi dalam nilai akhir. Evaluasi menjadi autopsy (bedah mayat) di akhir proses, bukan sebagai check-up (pemeriksaan kesehatan) yang berkelanjutan.

Fenomena dominasi tes sumatif dalam konteks pembelajaran bahasa asing bukan hal baru, banyak guru hanya menggunakan kuis, mid-term, dan final exam sebagai alat evaluasi formal, tanpa menyediakan umpan balik berbasis teks yang detail setelah tes, sehingga aspek proses pembelajaran tidak terefleksi secara memadai. Seringkali hasil tes angka atau huruf nilai diberikan

¹⁸ Hakim, M. Arif Rahman, and Andri Saputra. "Efek washback ujian nasional subyek bahasa Inggris pada tingkat sekolah menengah pertama (SMP) di Indonesia." *Tadbir: Jurnal Studi Manajemen Pendidikan* 4.1 (2020): 39-52.

tanpa penjelasan mendalam tentang kekuatan dan kelemahan siswa dalam berbagai keterampilan bahasa.¹⁹ Kondisi ini membatasi kesempatan bagi siswa untuk mengenali area yang butuh perbaikan, serta membatasi guru dalam menyesuaikan strategi pengajaran berdasarkan kebutuhan aktual mahasiswa.

Efek dari dominasi evaluasi sumatif terhadap motivasi, kecemasan, dan regulasi belajar mahasiswa juga telah diteliti, kelompok yang mendapatkan penilaian formatif menunjukkan peningkatan regulasi diri dan motivasi belajar dibanding yang hanya melalui ujian sumatif.²⁰ Penilaian sumatif terutama bila diposisikan sebagai penentu tunggal hasil belajar, cenderung menimbulkan kecemasan tinggi dan menghambat proses belajar berkelanjutan.

Hal tersebut menunjukkan bahwa dominasi evaluasi sumatif bisa berdampak negatif pada proses pembelajaran bahasa khususnya ketika tujuan pembelajaran adalah kemampuan komunikatif dan perkembangan berkelanjutan. Berikut tabel ringkasan dari beberapa penelitian terkait praktik dominasi evaluasi sumatif dan kelemahan minimnya asesmen formatif sebagai data pendukung:

Tabel 3. Penelitian Terdahulu Terkait Dominasi Evaluasi Sumatif dan Minimnya Asesmen Proses dalam Pembelajaran Bahasa

Penelitian/Studi	Konteks	Temuan Terkait Dominasi Sumatif/Kekurangan Asesmen Proses	Implikasi bagi Pengajaran Bahasa
Latif, Muhammad Wasim, and Arzoo Wasim (2024). ²¹	Pendidikan bahasa asing di sekolah menengah	Kuiz, mid-term, final exam dominan; umpan balik tekstual jarang diberikan setelah tes sumatif	Kemajuan siswa selama proses tidak termonitor; kemampuan produktif dan proses belajar kurang dievaluasi
Ismail, Seyed M., et al. (2022). ²²	Mahasiswa EFL di Iran	Kelompok dengan asesmen formatif menunjukkan motivasi dan regulasi belajar	Penilaian formatif mendukung perkembangan

¹⁹ Rachmawati, Ade Like, Aminnudin Aminnudin, and Kartimi Kartimi. "Optimalisasi Penilaian Berbasis Kelas dan Pengembangan Tes untuk Meningkatkan Literasi Siswa SDIT Wadi Fatimah Cirebon." *Ar-rrayan: Journal Of Islamic Education* 1.2 (2024): 102-115.

²⁰ Arjanggi, R., and E. A. Setiowati. "Meningkatkan belajar berdasar regulasi diri melalui pembelajaran kooperatif tipe jigsaw." *Makara Seri Sosial Humaniora* 17.1 (2013): 55-63.

²¹ Latif, Muhammad Wasim, and Arzoo Wasim. "Investigating EFL instructors' approaches to classroom-based assessment culture: An explanatory sequential mixed-method approach." *Language Testing in Asia* 14.1 (2024): 57.

²² Ismail, Seyed M., et al. "Formative vs. summative assessment: impacts on academic motivation, attitude toward learning, test anxiety, and self-regulation skill." *Language Testing in Asia* 12.1 (2022): 40.

		lebih baik, sedangkan fokus pada sumatif menimbulkan kecemasan	berkelanjutan, bukan hanya hasil akhir
Shafii, Rajabu Abdalah, and Jean-Louis Berger (2025). ²³	Pendidikan dasar/menengah di Tanzania	Sebagian besar guru masih mengandalkan ujian tertulis tradisional, sedikit menggunakan tugas autentik atau asesmen performa	Dominasi tes tradisional membatasi penilaian kompetensi nyata dan kemampuan komunikatif siswa
Yadi, Yogi Didi, et al. (2025). ²⁴	Perguruan tinggi di Indonesia	Tes sumatif dianggap stres bagi siswa; asesmen formatif membantu peningkatan berbicara dan rasa percaya diri	Untuk skill produktif seperti berbicara, perlu asesmen berbasis proses dan latihan berkelanjutan

Tabel ini memperkuat bahwa dominasi evaluasi sumatif bukan hanya masalah di Universitas Walailak, tetapi juga fenomena yang lebih luas di pengajaran bahasa asing dan pendidikan secara umum. Dalam BIPA di Walailak, dominasi evaluasi sumatif menyebabkan praktik belajar mengarah pada “lulus ujian” semata mahasiswa lebih fokus menghafal materi untuk menghadapi UTS/UAS daripada membangun kompetensi bahasa secara bertahap. Observasi di kelas menunjukkan bahwa meskipun tugas kecil, latihan berbicara, dan diskusi diberikan, karena tidak tercatat sebagai bagian dari nilai akhir secara sistematis, mahasiswa cenderung meremehkan atau kurang memprioritaskan aktivitas tersebut. Perasaan bahwa “nilai akhir menentukan segalanya” menciptakan tekanan dan menghambat motivasi intrinsik untuk terus belajar selama semester berlangsung.

Para pengajar dalam wawancara mengakui bahwa sistem penilaian yang ada membuat mereka jarang menilai proses misalnya partisipasi aktif, kemajuan berbicara, atau keterlibatan dalam diskusi karena beban administratif dan tuntutan mengoreksi UTS/UAS untuk banyak siswa dalam waktu singkat. Evaluasi formatif yang ideal sebagai alat untuk mendukung perkembangan tidak dilaksanakan secara konsisten. Situasi ini mendekati temuan dalam penelitian di Tanzania bahwa dominasi tes tradisional makin menguat meskipun kurikulum menekankan kompetensi.

Dominasi evaluasi sumatif juga mengurangi peluang untuk umpan balik yang bermanfaat:

²³ Shafii, Rajabu Abdalah, and Jean-Louis Berger. "Teacher assessment literacy, formative assessment practices, and their perceived efficacy in Tanzania: A scoping review." *Studies in Educational Evaluation* 86 (2025): 101496.

²⁴ Yadi, Yogi Didi, et al. "Analyzing the Effectiveness of Formative and Summative Assessments for EFL Speaking Skills: A Qualitative Study at Hamzanwadi University." *Journal of English Language and Education* 10.5 (2025): 929-939.

siswa tidak mendapat informasi detail tentang kelemahan mereka setelah tes, sehingga sulit bagi mereka untuk memperbaiki diri sebelum penilaian berikutnya. Tanpa umpan balik, proses pembelajaran bisa stagnan; siswa mungkin tidak menyadari area yang perlu dikembangkan, dan guru tidak memiliki data untuk menyesuaikan pengajaran berdasarkan kebutuhan siswa.²⁵

Program BIPA tingkat madya di Walailak berisiko menghasilkan lulusan yang secara “akademik” lulus karena nilai ujian memenuhi syarat namun secara fungsional kurang siap dalam penggunaan bahasa nyata, komunikasi lisan, dan pemahaman budaya, karena proses belajar dan pengembangan skill produktif tidak mendapat ruang evaluasi yang memadai. Skenario ini mereduksi tujuan program BIPA sebagai pembentuk kompetensi bahasa dan kultural bukan sekadar penghapalan materi untuk ujian.

Sangat disarankan agar program BIPA di Walailak melakukan reformasi dalam model evaluasi: mengurangi dominasi nilai ujian sumatif, memperkuat dan mensistematisasikan asesmen formatif berbasis proses sepanjang semester (misalnya observasi partisipasi, penilaian presentasi, tugas menulis & berbicara berkelanjutan, jurnal refleksi, proyek budaya), serta memberikan umpan balik informatif kepada mahasiswa. Evaluasi menjadi bagian integral dari proses pembelajaran dan membantu mahasiswa berkembang secara komprehensif bukan hanya mencapai nilai tinggi di akhir semester.

Tantangan Evaluasi Kompetensi Pragmatik dan Budaya

Materi tingkat madya mulai banyak menyisipkan unsur budaya dan pragmatik, seperti penggunaan ungkapan “memang” untuk penekanan, atau strategi menolak ajaran secara halus. Namun, mengevaluasi pemahaman ini sangat sulit dengan tes tertulis konvensional. Pengajar berusaha mengujinya melalui role-play, tetapi menghadapi kendala. Pertama, waktu yang terbatas untuk menilai semua mahasiswa. Kedua, subjektivitas penilaian. Ketiga, seperti hasil wawancara yang diungkapkan Responden P1:

“Mahasiswa Thailand sering malu dan kaku saat role-play menirukan percakapan orang Indonesia. Mereka takut salah dan tidak percaya diri. Ini bukan karena tidak paham materinya, tapi karena faktor budaya mereka yang sangat keng jai dan konteksnya asing.”

Evaluasi justru berpotensi mengukur tingkat kepercayaan diri dan adaptasi kultural, bukan kompetensi pragmatik semata. Evaluasi kompetensi pragmatik dan budaya memang sulit diukur dengan alat evaluasi tradisional. Meninjau berbagai instrumen (discourse completion tests, role-

²⁵ Damayanti, Ayu Maya, et al. *Evaluasi pembelajaran*. Basya Media Utama, 2023.

plays, self-assessment, digital assessment) dan mencatat bahwa setiap alat memiliki kelebihan dan keterbatasan tidak ada satu metode pun yang mampu mengukur pragmatik secara menyeluruh tanpa kompromi. Ini menunjukkan bahwa tantangan di BIPA Walailak bukan semata kelemahan lokal, melainkan bagian dari problem global dalam asesmen kompetensi pragmatik.

Bahkan ketika menggunakan Discourse Completion Tests (DCT) atau Multiple-Choice Discourse Completion Tests (MDCT), banyak siswa gagal menunjukkan kemampuan pragmatik yang memadai terutama dalam menyampaikan ungkapan kesopanan, penolakan halus, atau interaksi konteks sosial.²⁶ Walau teks atau konteks disajikan dalam soal, kemampuan untuk merealisasikan ungkapan pragmatik dalam interaksi nyata tetap sulit diukur sebuah masalah yang relevan dengan pengalaman mahasiswa BIPA di Thailand. Berikut tabel ringkasan beberapa penelitian empiris yang mengkaji evaluasi kompetensi pragmatik / budaya serta tantangan dalam mengukur kemampuan.

Tabel 4. Penelitian Terdahulu mengenai Instrumen Evaluasi Kompetensi Pragmatik dan Unsur Budaya dalam Pembelajaran Bahasa

Studi/Penelitian	Instrumen/Metode Evaluasi	Temuan/Tantangan dalam Mengukur Pragmatik/Budaya	Implikasi untuk Pembelajaran Bahasa Asing/BIPA
Paluanova, Khalifa (2024). ²⁷	Discourse Completion Test (DCT), Role-Play, Self-assessment, Digital tasks	Setiap alat memiliki kelebihan dan keterbatasan; tidak ada metode tunggal yang efektif mengukur pragmatik secara menyeluruh	Perlu kombinasi metode untuk validitas dan reliabilitas tinggi dalam menilai pragmatik
Bouknify, Mohammed (2025). ²⁸	Written & Multiple-Choice Discourse Completion Tests (WDCT / MDCT)	Banyak siswa gagal menunjukkan penguasaan kesopanan, penolakan halus, atau strategi komunikasi kontekstual	Tes tertulis saja tidak cukup; perlunya tugas performatif atau interaktif
Kusevska, Marija, et al. (2016). ²⁹	Role-play, DCT, analisis wacana	Menekankan bahwa role-play penting untuk	Penilaian pragmatik

²⁶ Devi, Sri, A. B. Yusuf, and M. Munirah. "Respons Terhadap Tuturan Sarkasme Ditinjau Dari Aspek Gender Dengan Discourse Completion Task (Dct)." *Kredo: Jurnal Ilmiah Bahasa dan Sastra* 5.2 (2022).

²⁷ Paluanova, Khalifa. "Evaluating Pragmatic Competence in Language Learners Through an Assessment of Current Measurement Tools." *Emergent: Journal of Educational Discoveries and Lifelong Learning (EJEDL)* 6.2 (2024).

²⁸ Bouknify, Mohammed. "Task-Based Assessment Of Pragmatic Competence: A Moroccan Efl High School Perspective." *International Online Journal of Education and Teaching (IOJET)*, 12(4), 2025, 389-407

²⁹ Kusevska, Marija, et al. "Profiling Pragmatic Competence of Foreign Language Learners." *Romanian Journal of English Studies* 13.1 (2016): 79-90.

		mengeksplorasi kemampuan pragmatik yang autentik, tetapi desain tugas, konteks dan instruksi harus dirancang dengan teliti agar valid	memerlukan konteks realistik dan rubrik penilaian jelas
Li, Rui, Rozina Raja, and Azimah Sazalie (2022). ³⁰	MDCT, WDCT, wawancara retrospektif, produksi wacana	Meskipun mahasiswa bisa mengenali konteks, banyak kesulitan dalam produksi bahasa yang pragmatis dan konteks sosial-budaya	Kesadaran pragmatik tidak otomatis diterjemahkan ke produksi pragmatik; butuh latihan berulang dalam konteks nyata

Tabel tersebut memperkuat bahwa pengukuran kompetensi pragmatik dan budaya rentan terhadap kelemahan metodologis: alat yang terlalu sederhana atau tidak kontekstual cenderung gagal mendeteksi kemampuan pragmatik, sementara metode interaktif atau performatif seperti role-play memerlukan desain dan rubrik yang cermat serta sumber daya hal yang sering menjadi kendala dalam praktik.

Kasus di Walailak kendala waktu dan jumlah mahasiswa menjadikan implementasi role-play secara menyeluruh sulit dilakukan, padahal peran konteks realistik dalam role-play sangat penting untuk mendapatkan data valid tentang pragmatik. Jika hanya sebagian mahasiswa yang berkesempatan, atau penilaian dilakukan secara subjektif tanpa rubrik komprehensif, maka hasil evaluasi bisa bias lebih mencerminkan keberanian atau kenyamanan budaya mahasiswa, dibandingkan penguasaan pragmatik.³¹

Pengajaran pragmatik secara eksplisit misalnya melalui tugas tertulis maupun lisan yang fokus pada ungkapan sopan, penolakan, permintaan, dan ekspresi sopan santun diperlukan agar siswa tidak hanya memahami teori, tetapi bisa mereproduksi secara tepat. jurnal.ar-raniry.ac.id+1 Tanpa pendekatan seperti ini, upaya evaluasi melalui role-play atau DCT saja akan gagal jika siswa belum dibekali awareness pragmatik sebelumnya.

Evaluasi kompetensi pragmatik dan budaya pada program BIPA tingkat madya di Universitas

³⁰ Li, Rui, Rozina Raja, and Azimah Sazalie. "An investigation into Chinese EFL learners' pragmatic competence." *GEMA Online Journal of Language Studies* 15.2 (2015): 101.

³¹ Musyrifa, Furaida Ayu, and Khabibi Muhammad Luthfi. "Metode Komunikatif dengan Role Play Berbasis Gender dalam Pembelajaran Kalām." *Alsina: Journal of Arabic Studies* 3.1 (2021): 103-136.

Walailak menghadapi tantangan sistemik: instrumen kurang memadai, waktu dan sumber daya terbatas, serta faktor budaya dan psikologis mahasiswa sebagai penutur asing. Evaluasi secara sumatif atau dengan soal tertulis saja jelas tidak mampu mengevaluasi aspek pragmatik dan kultural secara valid, sementara metode performatif seperti role-play harus dirancang dan dilaksanakan dengan serius agar mampu mencerminkan kompetensi pragmatik secara autentik.³²

urgensi reformasi model evaluasi dalam program BIPA, khususnya di konteks internasional seperti mahasiswa Thailand: perlu dikombinasikan evaluasi formatif dan formatif-performative, penggunaan rubrik pragmatik dan kultural yang transparan, serta pengajaran pragmatik secara eksplisit sebelum evaluasi. Evaluasi bukan hanya mengukur penguasaan grammar dan kosa kata, tetapi kemampuan nyata menggunakan bahasa dalam situasi kultural menjadikan lulusan BIPA benar-benar kompeten dalam konteks komunikasi nyata, bukan sekadar memenuhi target nilai.

Keterbatasan Kontekstualisasi Materi Evaluasi dengan Realitas Thailand

Bahan ajar yang berasal dari Indonesia seringkali menggunakan konteks yang sangat Indonesia-sentris (membahas transjakarta, gado-gado, atau sistem zonasi sekolah). Saat dievaluasi, mahasiswa dituntut memahami konteks asing tersebut. Hal ini menyulitkan karena beban kognitif mereka menjadi dobel: memahami bahasa sekaligus memahami konteks budaya asing. Pengajar sebenarnya telah berupaya mengkontekstualisasikan dengan contoh dari Thailand, tetapi belum merambah hingga ke level instrumen evaluasi. Soal-soal evaluasi masih banyak yang menggunakan konteks Indonesia. Seperti diungkapkan dalam jawaban wawancara dari seorang mahasiswa dalam FGD:

"Saya mengerti kata-katanya, tapi tidak mengerti mengapa masalahnya bisa seperti itu [dalam teks]. Kalau contohnya tentang songtaew [angkutan umum di Thailand] atau festival Loy Krathong, pasti lebih mudah."

Masalah ini sejalan dengan diskusi tentang "bias budaya dalam tes bahasa asing". Item-item tes kerap mengandung referensi budaya, kebiasaan, institusi, atau praktik sosial yang akrab bagi kalangan tertentu tetapi asing bagi kalangan lain sehingga performa peserta bisa dipengaruhi oleh latar budaya mereka, bukan kemampuan bahasa semata.³³ Ketidaksesuaian konteks budaya pada

³² Amil, Ahmad Jami'ul. "Internasionalisasi Bahasa Indonesia Melalui Study Abroad Program Bahasa Indonesia Bagi Penutur Asing (Bipa) Menggunakan Pendekatan Sastra Dan Seni (Studi Kasus Pada Mahasiswa Asean Studies Universitas Walailak Thailand Tahun 2020)." *Prosiding Seminar Internasional Seminar Kepakaran Bahasa Indonesia Untuk Penutur Asing (SEMAR BIPA)*. Vol. 3. No. 1. 2020.

³³ Djiwandono, Patrisius Istiarto. "Cultural bias in language testing." *TEFLIN Journal* 17.1 (2006): 81-89.

soal tes bisa membuat tes tersebut tidak adil dan tidak valid untuk semua peserta dengan latar budaya berbeda. Hal ini memperkuat argumentasi bahwa materi evaluasi BIPA yang terlalu Indonesia-sentris dapat merugikan mahasiswa non-Indonesia seperti mahasiswa Thailand di Walailak.

Dalam adaptasi asesmen kinerja/performance-based assessment pada konteks internasional ketika instrumen asesmen diterjemahkan atau diadaptasi ke bahasa dan budaya yang berbeda, sering muncul masalah equivalence kontekstual dan kultural: meskipun aspek linguistik diterjemahkan, konteks dan pengalaman yang mendasari soal bisa sulit dipindahkan secara setara.³⁴ Hal ini berarti bahwa adaptasi bahan ajar dari Indonesia ke mahasiswa Thailand bukan sekadar mengganti bahasa, tetapi perlu mempertimbangkan latar sosial-kultural dan konteks pengalaman agar instrumen evaluasi mengukur kompetensi sebenarnya, bukan kemampuan menafsirkan budaya asing.

Dalam program BIPA di Walailak akibat evaluasi yang menggunakan konteks budaya Indonesia bukan konteks Thailand atau lintas budaya mahasiswa Thailand menghadapi dua beban, pertama, menguasai bahasa. Kedua, memahami realitas sosial dan budaya asing. Sebagaimana diungkapkan dalam FGD, meskipun mahasiswa memahami kosakata dan struktur kalimat, mereka kadang kesulitan memahami konteks dan mengapa persoalan dalam teks bisa seperti itu sehingga kognisi linguistik mereka berjalan, tetapi interpretasi pragmatik budaya menjadi hambatan. Kondisi ini memang paralel dengan keluhan dalam kajian bias budaya: bahwa performa dalam tes bisa terganggu oleh unfamiliar context, bukan oleh rendahnya kemampuan bahasa.³⁵

Upaya beberapa pengajar untuk mengkontekstualisasikan materi misalnya memberi contoh yang relevan dengan realitas mahasiswa di Thailand menunjukkan bahwa ada kesadaran pedagogis atas pentingnya adaptasi konteks budaya. Adaptasi ini belum meluas ke instrumen evaluasi, soal-soal evaluasi tetap menggunakan referensi Indonesia, sehingga mismatch antara materi ajar, konteks mahasiswa, dan evaluasi tetap terjadi. Kondisi ini mengurangi validitas tes dan merusak tujuan pembelajaran bahasa sebagai alat komunikasi lintas budaya.

Hal lain yang perlu diperhatikan adalah bahwa adaptasi tes antar kebudayaan bukan hanya

³⁴ Ursin, Jani, et al. "Linguistic, Contextual, and Experiential Equivalence Issues in the Adaptation of a Performance-Based Assessment of Generic Skills in Higher Education." *Frontiers in education*. Vol. 7. Frontiers Media SA, 2022.

³⁵ Lanteigne, Betty. "Regionally Specific Tasks of Non-Western English Language Use." *TESL-EJ* 10.2 (2006): n2.

soal penggantian konteks, tetapi juga soal “*equivalence*” bahwa makna, fungsi sosial, pragmatik, dan budaya dalam teks atau tugas harus tetap setara agar hasil tes benar-benar mengukur kompetensi bahasa, bukan pengetahuan budaya asal. Faktor linguistik, kontekstual, dan pengalaman (experiential) berperan besar dalam bagaimana peserta memahami dan menanggapi tugas evaluasi.³⁶

Problem evaluasi dalam program BIPA di Universitas Walailak khususnya terkait kontekstualisasi materi dan realitas mahasiswa Thailand tidak bisa diabaikan. Evaluasi yang tetap menggunakan konteks asing bagi peserta berpotensi menghasilkan skor yang tidak valid, memunculkan bias budaya, serta mengukur adaptasi budaya dan kemampuan menafsir konteks asing, bukan kompetensi bahasa sejati.

KESIMPULAN

Berdasarkan analisis data, dapat disimpulkan bahwa evaluasi pembelajaran BIPA tingkat madya di Universitas Walailak menghadapi empat problematika inti yang saling berkaitan: (1) ketidakselarasan antara cakupan materi ajar dan instrumen evaluasi; (2) sistem evaluasi yang masih didominasi oleh pendekatan sumatif dan berorientasi produk; (3) kesulitan teknis dan metodologis dalam mengevaluasi kompetensi pragmatik dan budaya secara valid dan reliabel; serta (4) kurangnya kontekstualisasi materi evaluasi dengan realitas dan latar belakang pembelajar Thailand. Problematika ini tidak hanya berdampak pada keakuratan pengukuran kompetensi mahasiswa, tetapi juga pada motivasi belajar dan keadilan dalam penilaian.

REFERENSI

- Achjar, Komang Ayu Henny, et al. *Metode penelitian kualitatif: Panduan praktis untuk analisis data kualitatif dan studi kasus*. Jambi: PT. Sonpedia Publishing Indonesia, 2023.
- Alfiani, Calvin Reindira, Abdul Rani, and Sri Wahyuni. "Ketidaksesuaian Materi Dan Asesmen Pada Pembelajaran Bahasa Indonesia Teks Negosiasi." *Imajeri: Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 7 (2024).
- Amil, Ahmad Jami'ul. "Internasionalisasi Bahasa Indonesia Melalui Study Abroad Program Bahasa Indonesia Bagi Penutur Asing (Bipa) Menggunakan Pendekatan Sastra Dan Seni (Studi Kasus Pada Mahasiswa Asean Studies Universitas Walailak Thailand Tahun 2020)." *Prosiding Seminar Internasional Seminar Kepakaran Bahasa Indonesia Untuk Penutur Asing (SEMAR BIPA)*. Vol. 3. No. 1. 2020.

³⁶ Andidar, Denestyar Rizki, and Nurul F. Prahastuti. "Penerapan Konteks Budaya Dalam Psikometrika." *Jurnal Talenta Psikologi* 10.2 (2021).

- Andidar, Denestyar Rizki, and Nurul F. Prahastuti. "Penerapan Konteks Budaya Dalam Psikometrika." *Jurnal Talenta Psikologi* 10.2 (2021).
- Arjanggi, R., and E. A. Setiowati. "Meningkatkan belajar berdasar regulasi diri melalui pembelajaran kooperatif tipe jigsaw." *Makara Seri Sosial Humaniora* 17.1 (2013): 55-63.
- Bouknify, Mohammed. "Task-Based Assessment Of Pragmatic Competence: A Moroccan Efl High School Perspective." *International Online Journal of Education and Teaching (IOJET)*, 12(4), 2025, 389-407.
- Damawanti, Kustina Dwi Nur, Sri Wahyuni, and Frida Siswiyanti. "Implementasi Multimedia Interaktif Dalam Pembelajaran Menulis Cerita Pada Mahasiswa Asing Di Universitas Islam Malang." *Jurnal Penelitian, Pendidikan, dan Pembelajaran* 20.3 (2025).
- Damayanti, Ayu Maya, et al. *Evaluasi pembelajaran*. Basya Media Utama, 2023.
- Devi, Sri, A. B. Yusuf, and M. Munirah. "Respons Terhadap Tuturan Sarkasme Ditinjau Dari Aspek Gender Dengan Discourse Completion Task (Dct)." *Kredo: Jurnal Ilmiah Bahasa dan Sastra* 5.2 (2022).
- Djiwandono, Patrisius Istiarto. "Cultural bias in language testing." *TEFLIN Journal* 17.1 (2006): 81-89.
- Dong, Manxia. "Structural relationship between learners' perceptions of a test, learning practices, and learning outcomes: A study on the washback mechanism of a high-stakes test." *Studies in Educational Evaluation* 64 (2020): 100824.
- Hakim, M. Arif Rahman, and Andri Saputra. "Efek washback ujian nasional subyek bahasa Inggris pada tingkat sekolah menengah pertama (SMP) di Indonesia." *Tadbir: Jurnal Studi Manajemen Pendidikan* 4.1 (2020): 39-52.
- Haman Supriyadi et al. "*Harmoni Kata: Panduan Belajar Kosakata Indonesia-Thai*", Jakarta: SEAMEO QITEP in Language, 2023.
https://www.qiteplanguage.org/assets/files/dokumen/Harmoni%20Kata_Panduan%20Belajar%20Kosa%20Kata%20Bahasa%20Indonesia%20Thai-Copy.pdf.
- Ismail, Seyed M., et al. "Formative vs. summative assessment: impacts on academic motivation, attitude toward learning, test anxiety, and self-regulation skill." *Language Testing in Asia* 12.1 (2022): 40.
- Kusevska, Marija, et al. "Profiling Pragmatic Competence of Foreign Language Learners." *Romanian Journal of English Studies* 13.1 (2016): 79-90.
- Lanteigne, Betty. "Regionally Specific Tasks of Non-Western English Language Use." *TESL-EJ* 10.2 (2006): n2.
- Latif, Muhammad Wasim, and Arzoo Wasim. "Investigating EFL instructors' approaches to classroom-based assessment culture: An explanatory sequential mixed-method approach." *Language Testing in Asia* 14.1 (2024): 57.
- Li, Rui, Rozina Raja, and Azimah Sazalie. "An investigation into Chinese EFL learners' pragmatic competence." *GEMA Online Journal of Language Studies* 15.2 (2015): 101.
- Mintowati, Maria, et al. "Problematics Of BIPA Teachers At Unesa Indonesia And Walailak University Thailand." *International Conference of Humanities and Social Science (Ichss)*. 2021.

- Musyrifa, Furaida Ayu, and Khabibi Muhammad Luthfi. "Metode Komunikatif dengan Role Play Berbasis Gender dalam Pembelajaran Kalām." *Alsina: Journal of Arabic Studies* 3.1 (2021): 103-136.
- Paluanova, Khalifa. "Evaluating Pragmatic Competence in Language Learners Through an Assessment of Current Measurement Tools." *Emergent: Journal of Educational Discoveries and Lifelong Learning (EJEDL)* 6.2 (2024).
- Pramesti, Riza Dian. "Pengembangan Media Pembelajaran E-Modul (Book Creator) Aktivitas Liburan Menilik Pesona Indonesia Pada Pembelajaran BIPA Tingkat Dasar", *Skripsi*, Universitas Islam Malang, 2022.
- Rachmawati, Ade Like, Aminnudin Aminnudin, and Kartimi Kartimi. "Optimalisasi Penilaian Berbasis Kelas dan Pengembangan Tes untuk Meningkatkan Literasi Siswa SDIT Wadi Fatimah Cirebon." *Ar-rayyan: Journal Of Islamic Education* 1.2 (2024): 102-115.
- Septyani, Dwi Endah, Zainal Rafli, and Liliana Muliastuti. "Keterbacaan Wacana Buku Teks BIPA" Sahabatku Indonesia" Tingkat Madya." *Indonesian Language Education and Literature* 6.1 (2020): 13-23.
- Shafii, Rajabu Abdalah, and Jean-Louis Berger. "Teacher assessment literacy, formative assessment practices, and their perceived efficacy in Tanzania: A scoping review." *Studies in Educational Evaluation* 86 (2025): 101496.
- Sudarman, Silas, et al. *Evaluasi Pendidikan Abad 21: Inovasi, Spiritualitas, dan Relevansi*. Jambi: PT. Sonpedia Publishing Indonesia, 2025.
- Sumardi, Sumardi. "Performance-based assessment as a current trend in ELT: Investigating its washback effects on secondary-school students learning." *Kajian Linguistik dan Sastra* 2.1 (2017): 1-11.
- Suwarno, Suparjo Adi. "Kinerja kepala madrasah dalam melaksanakan fungsi manajemen pendidikan di MTs Alamiriyyah Blokagung Karangdoro Tegalsari Banyuwangi." *Innovative: Journal Of Social Science Research* 3.6 (2023): 7012-7029.
- Tiawati, Refa Lina. "Bahasa Indonesia di Thailand menjadi media diplomasi kebahasaan dan budaya di ASEAN melalui pengajaran BIPA." *Jurnal Gramatika: Jurnal Penelitian Pendidikan Bahasa dan Sastra Indonesia* 1.1 (2015): 29-44.
- Tonji, Nasrulah Ridwan, et al. "Analisis Setting Sosial dalam Penyusunan Bahan Ajar PAI di SMA Negeri 2 Limboto Kabupaten Gorontalo." *Konstruktivisme: Jurnal Pendidikan dan Pembelajaran* 16.2 (2024): 359-370.
- Ursin, Jani, et al. "Linguistic, Contextual, and Experiential Equivalence Issues in the Adaptation of a Performance-Based Assessment of Generic Skills in Higher Education." *Frontiers in education*. Vol. 7. Frontiers Media SA, 2022..
- Wibawani, Sri. *Pengembangan Instrumen Tes Berbasis Digital untuk Mengukur Kemampuan Membaca Mahasiswa BIPA Tingkat C1*. *Skripsi*. Universitas Brawijaya, 2021.
- Wigati, Prinastining Dyah, and Rudy Pramono. "EVALUASI PROGRAM PEMBELAJARAN BIPA DENGAN METODE CONTEXT, INPUT, PROCESS, PRODUCT (CIPP)." *Media Bina Ilmiah* 19.9 (2025): 5625-5638.

- Wulandari, Annisa Eka, and Neneng Konety. "Diplomasi Budaya Indonesia terhadap Vietnam melalui Program Bahasa Indonesia bagi Penutur Asing (BIPA)." *Padjadjaran Journal of International Relations* 6.2 (2024): 241-257.
- Yadi, Yogi Didi, et al. "Analyzing the Effectiveness of Formative and Summative Assessments for EFL Speaking Skills: A Qualitative Study at Hamzanwadi University." *Journal of English Language and Education* 10.5 (2025): 929-939.