

Fairness dalam Pengujian Kemahiran Reseptif Tes Evaluasi Belajar Bahasa Indonesia bagi Penutur Asing Tingkat I: Tinjauan berbasis Etik

Daud Yusuf^{1*}, Sisilia Setiawati Halimi²

^{1,2} Universitas Indonesia

e-mail: ysufdaud@gmail.com¹, halimisisilia@gmail.com²

Abstrak

Tes Evaluasi Pembelajaran Bahasa Indonesia Penutur Asing Tingkat I (TEB BIPA I) merupakan ujian untuk menilai prestasi belajar peserta didik dalam mempelajari bahasa Indonesia sebagai bahasa asing. Sebagai prototipe, TEB BIPA I membutuhkan evaluasi yang komprehensif. Tinjauan ini berusaha untuk menguji kegunaan dan kewajaran TEB BIPA I secara keseluruhan menggunakan model terbaru Kunnan (2018) untuk menilai kewajaran tes di bawah pendekatan berbasis etika. Tinjauan tersebut menggunakan tes keterampilan reseptif sebagai unit analisis. Data yang dikumpulkan meliputi studi dokumen naskah tes, spesifikasi tes, referensi kurikulum, hasil tes percontohan, halaman web BIPA Online dan petunjuk teknis pengembangan tes. Data dianalisis dengan menggunakan kerangka keadilan pendekatan berbasis etika Kunnan yang menganalisis klaim (1) kesempatan belajar, (2) kebermaknaan tes, (3) tidak adanya bias, dan (4) aksesibilitas.

Kata kunci: *Review Tes, BIPA, Evaluasi, Pendekatan Berbasis Etika*

Abstract

Tes Evaluasi Belajar Bahasa Indonesia bagi Penutur Asing Tingkat I (TEB BIPA I) is the exam for assessing learners' achievement in learning Indonesian as a foreign language. As a prototype, TEB BIPA I needs a comprehensive evaluation. This review seeks to examine the overall usefulness and fairness of TEB BIPA I using Kunnan's (2018) recent model of assessing test fairness under an ethical-based approach. The review took a receptive skills test as an analysis unit. The data collected involved a document study of the test script, test specifications, curriculum reference, pilot test results, BIPA Daring webpage and technical instructions for test development. The data thus analyzed using the fairness framework of Kunnan's ethical based approach which analyses the claim of (1) learning opportunity, (2) test meaningfulness, (3) the absence of bias, and (4) accessibility.

Keywords : *Test Review, BIPA, Evaluation, Ethical-Based Approach*

PENDAHULUAN

Tes merupakan pembuka jalan dalam pengambilan keputusan secara tidak acak dan adil bagi penggunaannya (Fulcher, 2010, hlm. 4). Dalam ranah kebahasaan, Brown dan Abeywickrama (2019) menyebut tes sebagai "metode yang digunakan untuk mengukur kemahiran, pengetahuan, dan kinerja seseorang pada bidang tertentu" (hlm. 3). Para ahli menggolongkan tes berdasarkan tujuannya ke dalam lima macam, yaitu tes pencapaian akademik, tes kemahiran berbahasa, tes diagnostik, tes bakat dan tes penempatan tingkat (Brown & Abeywickrama, 2019; Carr, 2011; Cheng & Fox, 2017; Hughes, 2003).

Salah satu tes capaian akademik. dalam konteks bahasa Indonesia bagi penutur asing (BIPA) adalah Tes Evaluasi Belajar Bahasa Indonesia bagi Penutur Asing (TEB BIPA). TEB BIPA merupakan ancangan tes standar dalam mengukur capaian peserta program belajar BIPA (Badan Pengembangan dan Pembinaan Bahasa, 2020) TEB BIPA disusun dan dikembangkan Pusat Penguatan dan Pemberdayaan Bahasa, Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. Jumlah

level pada TEB BIPA sesuai dengan jumlah level pembelajaran BIPA yang direkomendasikan kementerian yaitu tujuh level (Kemendikbud, 2016; 2017). Pelevelan BIPA didasarkan pada Permendikbud Nomor 70 tahun 2016 tentang Standar Kemahiran Berbahasa Indonesia dan diselaraskan dengan penutur asing pada Permendikbud Nomor 27 tahun 2017 tentang Standar Kompetensi Lulusan Bahasa Indonesia bagi Penutur Asing. Saat ini, TEB BIPA masih dalam tahap pengembangan. TEB BIPA dikembangkan sejak tahun 2016 dan belum diluncurkan hingga saat ini.

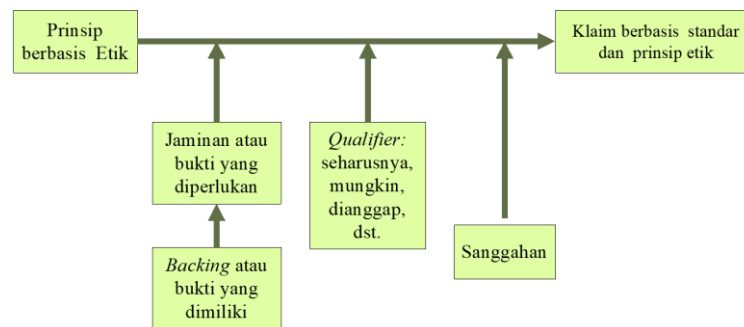
Sebagai tes yang berupa purwarupa dan dalam pengembangan, hingga saat ini belum ada tinjauan kritis dari sudut pandang pengembang dan pemangku kepentingan mengenai TEB BIPA. Tulisan ini mengangkat tinjauan kritis dari berbagai sudut pandang baik dari segi pengembang maupun pemangku kepentingan dengan mengevaluasi tes berdasarkan kerangka tinjauan yang digagas oleh ahli. Salah satu kerangka tinjauan yang sering digunakan dalam meninjau kritis tes kebahasaan adalah kerangka kebermanfaatan tes Bachman & Palmer (1996) seperti pada beberapa publikasi mengenai evaluasi tes yaitu tinjauan DELF (Elder, 2018) sebagai tes kemahiran bahasa Prancis, TestDaF (Norris & Drackert, 2018) sebagai tes kemahiran bahasa Jerman dan DuoLingo (Wagner, 2020) sebagai tes kemahiran bahasa Inggris. Tinjauan tes tersebut menggunakan kerangka Bachman dan Palmer (1996) yang membahas mengenai aspek kebermanfaatan validitas, reliabilitas, autentisitas, aksesibilitas, keadilan/integritas, dan dampak. Sejalan dengan kerangka kebermanfaatan Bachman dan Palmer (1996), Moghadam dan Nasirzadeh (2020) menulis tinjauan mengenai tes membaca bagi pemelajar bahasa Inggris dengan menerapkan kerangka *fairness* tes Kunnan (2004). Kerangka *fairness* tes Kunnan menitikberatkan tinjauan pada aspek validitas, bias, kemudahan akses, penyelenggaraan tes, dan konsekuensi sosial. Tinjauan tes kebahasaan kemudian beralih pada kerangka validasi berbasis argumen (Kane, 2004) seperti pada tinjauan tes bahasa Belanda ITNA/STRT (Deygers, Van den Branden, & Van Gorp, 2018) dan kerangka argumen penggunaan asesmen (Bachman & Palmer, 2010) seperti yang digunakan Peng, dkk. (2021) dalam meninjau tes bahasa Mandarin HSK. Kunnan (2018) mengembangkan pendekatan berbasis etik sebagai pengembangan dari kerangka Bachman & Palmer (1996), Kunnan (2004), Kane (2004), dan Bachman & Palmer (2010). Masih terbatas penelitian yang menggunakan pendekatan tersebut, khususnya dalam konteks tes bahasa Indonesia bagi penutur asing. Salah satu penerapan kerangka berbasis etik adalah tinjauan Celpe-Bras sebagai tes kemahiran berbahasa Portugis-Brasilia oleh Zhao & Liu (2019).

Tulisan ini menerapkan pendekatan berbasis etik (Kunnan, 2018; 2020) yang cukup baru dalam meninjau asesmen kebahasaan. Isu etik dianggap penting dalam tes bahasa. Fulcher (2010) sendiri menggagas istilah *adil* dalam definisinya mengenai tes. Isu etik mencakup *fairness*, *justice*, dan nilai menjadi arah baru dalam mengevaluasi tes, khususnya dalam tes bahasa seperti pada paparan Davies (1997) yang menggagas pembahasan kebajikan (*virtue*) tes tentang aspek moral dalam penyelenggaraan tes, Kunnan (1997) mengenai *fairness* dalam hal validasi nilai, Kunnan (2018) mengenai kesempatan belajar dalam mempersiapkan tes, Kunnan (2012) mengenai konsekuensi dan manfaat, dan Liu dan Kunnan (2015) serta Hoang dan Kunnan (2016) mengenai umpan balik.

Pendekatan berbasis etik merujuk pada “perspektif dari dunia filsafat moral dan pengetahuan etika atau etika dapat digunakan untuk membenarkan praktik individu dan institusi secara moral” (Kunnan, 2020, hlm. 84). Pendekatan berbasis etik mempunyai dua prinsip yang berlandaskan filosofi etik deontologis Immanuel Kant yaitu prinsip *fairness* dan prinsip *justice* (Kunnan, 2018). Selanjutnya, Kunnan (2020) mendeskripsikan prinsip *fairness* dengan gagasan bahwa “suatu asesmen seharusnya setara untuk semua peuji, yakni adanya praduga untuk memperlakukan setiap peuji setara” dan prinsip *justice* dengan deskripsi bahwa ‘suatu asesmen seharusnya adil, membawa manfaat dalam masyarakat, mendorong nilai-nilai positif, dan mendahulukan keadilan melalui justifikasi publik dan pemikiran” (hlm. 86).

Dalam penerapannya, Kunnan (2018) menggunakan baik pendekatan berbasis standar maupun pendekatan berbasis argumen dan mengembangkan pendekatan tersebut dengan

menambahkan landasan etik pada tinjauan mengenai asesmen bahasa. Kunnan (2018) menggunakan diagram Toulmin (2003) untuk “mengakomodasi prinsip-prinsip etik yang mendasari pengartikulasian klaim” (hlm. 86) seperti diilustrasikan pada gambar sebagai berikut.



Gambar 1 Model Toulmin berbasis etik dalam evaluasi asesmen (Kunnan, 2020, hlm. 87)

Tulisan ini meninjau seksi pengujian kemahiran reseptif dengan menerapkan prinsip *fairness* dalam tinjauan berbasis etik (Kunnan, 2018). Prinsip *fairness* mengacu pada keadilan dalam asesmen dalam sudut pandang peuji. Secara spesifik, prinsip berbasis etik mengarahkan pada klaim umum dan dikembangkan lagi pada beberapa sub-klaim. Klaim dan sub-klaim yang akan dibuktikan dengan alur seperti pada Gambar 1, yaitu pembahasan mengenai bukti yang diperlukan dan *backing* yang relevan, menentukan *qualifier* atau tingkat keyakinan dari klaim yang diberikan, memberikan sanggahan dari klaim dan bukti, serta memberikan tinjauan berdasarkan klaim yang diajukan. Prinsip *fairness* dalam tinjauan ini dideskripsikan bahwa “Seksi Pengujian Kemahiran Reseptif pada Tes Evaluasi Belajar Bahasa Indonesia bagi Penutur Asing Tingkat I (TEB BIPA I) seharusnya adil (*fair*) bagi semua peuji dengan praanggapan peuji telah diberikan kesempatan yang sama untuk mendemonstrasikan kemahirannya”. Berikut deskripsi mengenai klaim dari prinsip *fairness* dengan aspek kesempatan belajar, kebermaknaan, ketiadaan bias, dan aksesibilitas.

Tabel 1 Deskripsi Klaim Prinsip *Fairness*

Klaim	Deskripsi
Klaim: Prinsip <i>Fairness</i>	Seksi pengujian kemahiran reseptif TEB BIPA I adil untuk semua peuji
Klaim 1: Kesempatan Belajar	Sebelum mengambil pengujian kemahiran reseptif TEB BIPA I, terdapat kesempatan belajar dan persiapan yang memadai.
Klaim 2: Kebermaknaan	Seksi pengujian kemahiran reseptif TEB BIPA I bermakna dan konsisten.
Klaim 3: Ketiadaan Bias	Seksi pengujian kemahiran reseptif TEB BIPA I terbebas dari bias.
Klaim 4: Aksesibilitas	Seksi pengujian kemahiran reseptif TEB BIPA I menggunakan akses dan penyelenggaraan yang layak.

(diadaptasi dari Kunnan, 2020)

Berdasarkan paparan tersebut, tulisan ini mencoba untuk meninjau seksi pengujian reseptif TEB BIPA I melalui perspektif pendekatan berbasis etik dengan mencoba menjawab pertanyaan, “Bagaimana penerapan prinsip *fairness* dalam pendekatan berbasis etik pada seksi pengujian reseptif TEB BIPA I?”.

METODE PENELITIAN

Studi mengenai tinjauan tes termasuk dalam analisis konten kualitatif (Kunnan, 2018). Analisis konten kualitatif merupakan metode yang secara sistematis mendeskripsikan makna

dari data kualitatif (Flick, 2014). Penelitian ini menggunakan desain evaluasi pragmatik-partisipatoris (Dahler-Larsen, 2018) yaitu sebagai penelitian evaluasi yang mempertimbangkan berbagai sudut pandang dari pemangku kepentingan terlibat, serta sudut pandang tersebut dijadikan sebagai kriteria dan proses evaluasi itu sendiri. Penelitian mengenai evaluasi melibatkan empat aspek yaitu *evaluand* (objek yang dievaluasi), nilai, peran evaluasi, dan pendekatan yang hati-hati, sistematis, dan metodologis dalam menciptakan pengetahuan (Dahler-Larsen, 2018). *Evaluand* dalam konteks penelitian ini adalah seksi pengujian kemahiran reseptif pada Tes Evaluasi Belajar Bahasa Indonesia bagi Penutur Asing Tingkat I (TEB BIPA I). Nilai yang diusung dalam tinjauan ini adalah *fairness* dalam pengembangan tes sesuai dengan prinsip pendekatan berbasis etik (Kunnan, 2018). Evaluasi ini memberikan gambaran umum mengenai tinjauan secara kritis baik dalam sudut pandang pengembangan maupun penggunaannya, serta mendaftar agenda penelitian yang diperlukan dalam memberikan *evidence/backing* untuk klaim dalam penggunaan tes. Pendekatan berbasis etik (Kunnan, 2018) dipilih sebagai kerangka yang digunakan dalam meninjau secara kritis TEB BIPA I dan dianggap sebagai pendekatan yang metodologis dan sistematis (Kunnan, 2020) untuk meninjau tes, khususnya dalam konteks tes kebahasaan.

Pengumpulan data dilakukan melalui studi dokumentasi. Studi dokumentasi merupakan studi yang mengumpulkan dan menganalisis dokumen-dokumen, catatan-catatan yang penting dan berhubungan serta dapat memberikan data-data untuk memecahkan permasalahan dalam penelitian (Creswell & Creswell, 2018). Enam jenis dokumen digunakan sebagai instrumen dalam penelitian ini, yaitu naskah Tes Evaluasi Belajar (TEB) BIPA I, Kisi-kisi TEB BIPA I, Kurikulum acuan BIPA, Hasil uji coba TEB BIPA I, Laman BIPA Daring (bipa.kemdikbud.go.id), dan Petunjuk teknis penyusunan Tes Evaluasi Belajar BIPA.

Dalamancangan kualitatif, analisis data melibatkan tiga tahapan yaitu kondensasi data, *display* data, dan penarikan kesimpulan serta verifikasi data (Miles, Huberman, & Saldaña, 2014). Selain itu, secara spesifik analisis data dalam studi ini melibatkan analisis isi dokumen dengan tahapan analisis menggunakan analisis penerapan prinsip *fairness* dalam pendekatan berbasis etik (Kunnan, 2018) dengan menganalisis klaim mengenai (1) kesempatan belajar, (2) kebermaknaan, (3) ketiadaan bias, dan (4) aksesibilitas.

HASIL DAN PEMBAHASAN

Deskripsi Umum Tes

TEB BIPA I merupakan tes evaluasi belajar atau tes capaian akademik (*achievement*). Tes ini bertujuan sebagai alat ukur pengetahuan dan keterampilan peserta BIPA setelah menyelesaikan level tertentu, dalam konteks tinjauan ini adalah Level 1, dalam waktu tertentu berdasarkan standar kompetensi lulusan (SKL) BIPA yang telah ditetapkan dalam Peraturan Menteri Pendidikan dan Kebudayaan Nomor 27 Tahun 2017 tentang Standar Kompetensi Lulusan Kursus dan Pelatihan Bahasa Indonesia bagi Penutur Asing (Badan Pengembangan dan Pembinaan Bahasa, 2020).

TEB BIPA 1 terdiri atas empat bagian dengan jumlah soal keseluruhan 96 soal. Bagian pertama merupakan tes kemahiran menyimak dengan jumlah soal 40 soal. Selanjutnya, bagian kedua merupakan tes kemahiran membaca dengan jumlah soal 50 soal. Bagian ketiga merupakan tes kemahiran berbicara dengan jumlah soal 3 soal dan bagian terakhir merupakan tes kemahiran menulis dengan jumlah soal 3 soal. Peuji hanya diminta mengerjakan satu soal saja (memilih satu dari tiga pilihan soal) pada bagian tes kemahiran berbicara dan menulis. Bentuk soal pada tes kemahiran menyimak dan membaca adalah pilihan ganda dengan empat pilihan jawaban, sedangkan bentuk soal pada tes kemahiran berbicara dan menulis berupa uraian/kinerja.

Batas waktu pengerjaan soal seluruhnya adalah 120 menit. Batas waktu yang diberikan untuk mengerjakan 40 soal menyimak adalah 35 menit (sudah termasuk instruksi). Batas waktu yang diberikan untuk mengerjakan 50 soal membaca adalah 55 menit. Batas waktu untuk mengerjakan satu soal menulis yang berupa uraian/unjuk kerja adalah 25 menit, serta batas waktu untuk mengerjakan satu soal berbicara yang berupa unjuk kerja adalah 5 menit. TEB BIPA 1 (edisi 2019) diselenggarakan secara luring menggunakan uji tertulis (*paper-*

based) sementara mulai 2021, Badan Bahasa melalui Pustanda tengah mengembangkan Tes Evaluasi Belajar Bahasa Indonesia bagi Penutur Asing dalam Jaringan.

Pada uji kemahiran menyimak dan membaca, setiap jawaban benar mendapat bobot skor 1 dan jawaban salah mendapat bobot skor 0. Skor total pada bagian menyimak dan membaca dikonversikan dalam persentase sehingga total skor adalah jumlah jawaban benar dari jumlah soal dikalikan 100% dengan rentang skor 0–100. Untuk bagian berbicara dan menulis, TEB BIPA 1 mengadaptasi kriteria penilaian mengacu pada rubrik yang telah ditetapkan (Badan Pengembangan dan Pembinaan Bahasa, 2020). Jumlah skor maksimal keseluruhan dari bagian menulis dan berbicara adalah masing-masing 100 sehingga total skor keseluruhan dari empat bagian adalah 400. Akan tetapi, predikat penilaian diberikan untuk tiap-tiap bagian dengan rentang nilai sebagai berikut.

Tabel 3 Predikat Penilaian

Rentang Nilai	Predikat	Rentang Nilai	Predikat
96–100	Sempurna	46–55	Hampir Sedang
86–95	Baik Sekali	36–45	Kurang
76–85	Baik	26–35	Kurang Sekali
66–75	Cukup	16–25	Buruk
56–65	Sedang	0–15	Buruk Sekali

(diadaptasi dari Badan Pengembangan dan Pembinaan Bahasa, 2020)

Merujuk pada SKL BIPA, peuji yang lulus dalam TEB BIPA 1 mempunyai capaian kompetensi “mampu memahami dan menggunakan ungkapan konteks perkenalan diri dan pemenuhan kebutuhan konkret sehari-hari dan rutin dengan cara sederhana untuk berkomunikasi dengan mitra tutur yang sangat kooperatif” dengan deskripsi capaian pembelajaran khusus sebagai berikut.

Tabel 4 Deskripsi Capaian Pembelajaran Khusus

Aspek	Deskriptor
Kinerja	Memahami dan menggunakan ungkapan konteks perkenalan diri dan pemenuhan kebutuhan konkret sehari-hari dan rutin dengan cara sederhana untuk berkomunikasi dengan mitra tutur yang sangat kooperatif.
Pengetahuan	<p>Menguasai penggunaan tata bahasa dan kosa kata dalam berbagai jenis teks yang diajarkan yang meliputi.</p> <ol style="list-style-type: none"> 1. Penggunaan bunyi bahasa dan pelafalan. 2. Penggunaan ganti orang. 3. Penggunaan struktur frasa benda (DM). 4. Penggunaan kata bilangan tingkat. 5. Penggunaan kata negasi. 6. Penggunaan kalimat sederhana. 7. Penggunaan kata tanya. 8. penggunaan kata ganti tunjuk. 9. penggunaan kata kerja ada. 10. Penggunaan posisi dan Lokasi. 11. Penggunaan kata depan. 12. Penggunaan kata kerja berimbuhan. 13. Penggunaan kata keterangan. 14. Penggunaan kata hubung. 15. Penggunaan ungkapan dan kata sapaan, dan 16. Penggunaan kosa kata yang berhubungan dengan topik umum.

Diambil dari Kementerian Pendidikan dan Kebudayaan, 2017

Penyelenggaraan TEB BIPA 1 hingga saat ini masih bersifat terbatas dan sesuai permintaan dengan harga yang belum ditetapkan. Kontak pengembang TEB BIPA adalah sebagai berikut.

Pusat Penguatan dan Pemberdayaan Bahasa
Badan Pengembangan dan Pembinaan Bahasa
Kawasan IPSC Jl. Anyar Km. 4 Sukahati Citeureup Kab. Bogor 16810
Telepon: (021) 29099245; Posel: pustanda@kemdikbud.go.id

Penerapan Prinsip *Fairness* dalam Uji Kemahiran Reseptif TEB BIPA I

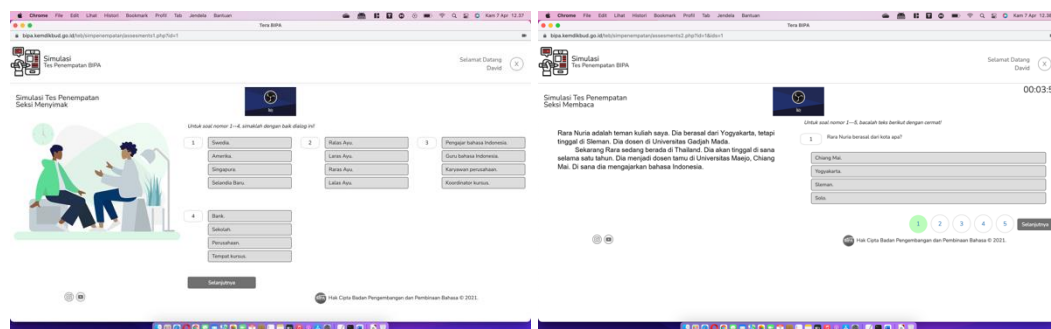
Prinsip *fairness* dalam tinjauan ini diartikulasikan bahwa “Seksi Pengujian Kemahiran pada Tes Evaluasi Belajar Bahasa Indonesia bagi Penutur Asing Tingkat I (TEB BIPA I) seharusnya adil (*fair*) bagi semua peuji dengan praanggapan peuji telah diberikan kesempatan yang sama untuk mendemonstrasikan kemahirannya”. Untuk mengargumentasikan klaim tersebut, bagian ini akan memaparkan deskripsi dari klaim prinsip *fairness* pada TEB BIPA I dengan membahas sub-klaim dari empat aspek yaitu kesempatan belajar, kebermaknaan, ketiadaan bias, dan aksesibilitas

1. *Sebelum mengambil seksi pengujian kemahiran reseptif pada TEB BIPA I, tersedia kesempatan belajar dan persiapan yang memadai*

Dalam memenuhi klaim ini, TEB BIPA 1 perlu memenuhi empat kriteria mengenai kesempatan belajar, yaitu tersedianya (1) kesempatan belajar yang memadai, (2) waktu persiapan tes yang memadai, (3) latihan dengan teknologi terkini, dan (4) kesempatan yang memadai tersedia dalam domain semiotik dan praktik sosial yang relevan. .

Untuk memenuhi kriteria mengenai tersedianya kesempatan belajar yang memadai, sebelum mengambil uji kemahiran reseptif TEB BIPA 1, kesempatan yang memadai (dalam hal kurikulum, materi, dan umpan balik) untuk memperoleh pengetahuan, kemampuan, atau keterampilan yang akan diuji telah tersedia. TEB BIPA merupakan tes evaluasi belajar yang mengacu pada kurikulum terstandar BIPA yang dikembangkan oleh Badan Pengembangan dan Pembinaan Bahasa. Kurikulum yang disediakan juga sesuai dengan materi yang diujikan serta program pembelajaran yang ditawarkan. Secara detail, kesesuaian antara materi uji dengan kurikulum dibahas pada bagian kebermaknaan (sub-klaim 2). Dalam hal materi pembelajaran, TEB BIPA mengacu pada bahan ajar yang dikembangkan oleh Badan Bahasa yang tersedia secara daring pada laman BIPA daring. Selain itu, saat ini Badan Bahasa bekerja sama dengan Perwakilan Republik Indonesia di berbagai negara menyelenggarakan kursus Bahasa Indonesia bagi Penutur Asing baik secara luring maupun secara daring untuk membuka akses yang seluas-luasnya dalam memfasilitasi pembelajaran bahasa Indonesia. Badan Bahasa saat ini sedang merancang sistem manajemen pembelajaran (*learning management system*) mandiri secara daring melalui platform BIPA Daring untuk memudahkan pemelajar dalam mengambil kelas bahasa Indonesia di mana saja dengan sumber pembelajaran tersedia di laman BIPA Daring tersebut.

Selanjutnya, dalam memenuhi kriteria mengenai ketersediaan waktu persiapan sebelum mengambil tes, sebagai tes pencapaian, TEB BIPA sudah tentu memberikan waktu persiapan yang cukup dalam hal persiapannya. Peuji yang akan mengambil TEB BIPA merupakan pemelajar BIPA yang mengambil program pembelajaran bahasa Indonesia sebelumnya dengan jumlah jam minimal yang diambil 180 jam pembelajaran pada tingkat BIPA 1 (Badan Pengembangan dan Pembinaan Bahasa, 2020). Kriteria selanjutnya adalah mengenai latihan dengan teknologi terkini. TEB BIPA memberikan fasilitas simulasi bagi para calon peuji sebelum mengambil tes melalui laman BIPA Daring (<https://bipa.kemdikbud.go.id/tera>). Simulasi yang disediakan berupa latihan soal seperti pada cuplikan layar berikut.



Gambar 4.1 Cuplikan Layar Simulasi TEB (Hak Cipta Badan Pengembangan dan Pembinaan Bahasa, 2022)

Kriteria terakhir sebagai aspek dari klaim mengenai kesempatan belajar adalah tersedianya kesempatan yang memadai tersedia dalam domain semiotik dan praktik sosial yang relevan. Seperti yang dipaparkan Gee (2003), domain semiotik didefinisikan sebagai kesadaran kolektif yang dimiliki oleh orang-orang dengan minat, atribut atau keahlian yang sama. Domain semiotik baik di dalam maupun di luar kelas dalam pembelajaran sangat diperlukan dalam hal pemelajaran yang aktif dan kritis sebagai pemanfaatan penuh sumber daya makna verbal dan non-verbal. Dalam hal ini, pembelajaran di dalam kelas tidak hanya semata mengajarkan buku teks tetapi juga menciptakan makna dari sumber daya semiotik mengenai Indonesia. Selain itu, komunitas Indonesia, forum pembelajaran bahasa Indonesia, video *youtube* mengenai pembelajaran juga menjadi salah satu sarana dalam mewujudkan pengalaman semiotik tersebut. Laman BIPA Daring juga diupayakan dapat mewujudkan pengalaman tersebut dalam satu wadah yang terintegrasi.

Selain itu, untuk memenuhi klaim mengenai kesempatan belajar yang cukup dalam mempersiapkan tes, pengembang tes perlu menyediakan informasi yang lengkap dan rinci dalam media yang dapat diakses oleh calon peuji serta menyediakan panduan bagi peuji yang berisi mengenai informasi mengenai tes dan latihan/simulasi sebelum mengambil TEB BIPA 1, seperti yang dipaparkan Zhao dan Liu (2019) bahwa dalam memenuhi klaim mengenai kesempatan belajar, informasi mengenai tes dan dokumen-dokumen berkaitan dengan persiapan tes: tes aktual yang diselenggarakan tahun ke tahun lengkap dengan video dan audio, panduan untuk peuji yang berisi contoh tes, jawaban, panduan pengisian serta komentar dari penyelenggara, serta panduan lain yang berisi informasi mengenai alur pendaftaran dan struktur dari penyelenggaraan tes.

Seperti yang telah dipaparkan sebelumnya, sanggahan dalam aspek kesempatan belajar ini berada pada latihan yang disediakan oleh pengembang. Latihan yang disediakan saat ini masih berupa purwarupa belum dapat digunakan sebagaimana mestinya. Selanjutnya, dalam menyediakan pengalaman semiotika, laman BIPA Daring diupayakan menjadi wadah terintegrasi yang dapat mewujudkan pengalaman tersebut. Selain itu, publikasi mengenai tes yang memadai juga perlu menjadi catatan dalam memenuhi kriteria mengenai kesempatan belajar. Publikasi yang disediakan perlu mencakup informasi umum, buku petunjuk untuk calon peuji, latihan soal yang diambil dari tes aktual beserta media pendukung yang memadai (Zhao dan Liu, 2019). Zhao dan Liu juga menggarisbawahi mengenai kesesuaian antara kompleksitas bahasa dari informasi yang disediakan dengan level kemahiran dari peuji. Penggunaan dwibahasa dengan bahasa jati peuji dalam informasi mengenai tes sangat direkomendasikan.

Berdasarkan paparan klaim mengenai kesempatan belajar dan persiapan yang memadai, TEB BIPA 1 dapat dikatakan memenuhi sebagian kriteria mengenai kesempatan belajar yang memadai, waktu persiapan tes yang memadai, latihan dengan teknologi terkini, serta perwujudan pengalaman semiotika bagi calon peuji. Hal yang menjadi catatan bahwa dalam memenuhi klaim mengenai kesempatan belajar, penyelenggara tes perlu memberikan informasi yang lengkap dan rinci kepada calon peuji, serta menyusun panduan pelaksanaan tes yang komprehensif bagi calon peuji.

2. Seksi pengujian kemahiran reseptif pada TEB BIPA I bermakna dan konsisten

Dalam memenuhi klaim ini, uji kemahiran reseptif pada TEB BIPA 1 perlu memenuhi tujuh kriteria mengenai kebermaknaan, yaitu (1) bermakna dalam hal pengembangan dan spesifikasi sesuai dengan tujuan kurikulum, (2) bermakna dalam hal konstruk dari kemahiran berbahasa yang diuji, (3) bermakna dalam hal variasi kebahasaan, konten, dan topik uji, (4) bermakna dalam hal mampu memprediksi kemahiran dengan kriteria eksternal, (5) konsisten dalam hal set butir/tugas dari konstruk yang berbeda, (6) konsisten dalam tugas uji yang beragam, serta (7) konsisten dengan beragam penguji yang terlibat dalam asesmen.

Kriteria pertama dalam aspek kebermaknaan adalah bahwa uji kemahiran reseptif pada TEB BIPA 1 bermakna dalam hal pengembangan dan spesifikasi sesuai dengan tujuan kurikulum melalui hasil analisis kesesuaian antara kisi-kisi, butir soal dengan SKL BIPA, bagian kemahiran reseptif pada TEB BIPA 1 (Chapelle, 2021). Hasil analisis kesesuaian antara kisi-kisi, butir soal dengan SKL BIPA sebagai tujuan kurikulum menunjukkan kecocokan yang tinggi antara butir soal dengan tujuan kurikulum. Kisi-kisi mendeskripsikan secara detail kompetensi dasar yang diuji dan diturunkan melalui indikator. Dalam kisi-kisi juga telah dicantumkan materi yang disajikan, rambu-rambu penyusunan soal, serta tingkat kesulitan soal.

Selain itu, seksi pengujian kemahiran reseptif pada TEB BIPA 1 dianggap bermakna ketika memenuhi kriteria kedua, yaitu bermakna dalam hal konstruk dari kemahiran berbahasa yang diuji. Melalui kerangka analisis ranah penggunaan bahasa target (Bachman dan Palmer, 1996), tes memang tidak sepenuhnya merepresentasikan situasi target dalam kehidupan nyata, akan tetapi uji reseptif TEB BIPA I telah mewakili situasi tersebut. Seperti dalam hal karakteristik input, kanal input berupa media cetak dan audio simakan mewakili media cetak/cetak digital dan audiovisual dalam situasi target. Bentuk teks verbal dalam tes juga mewakili teks multimodal yang ada dalam situasi. Karakteristik bahasa yang diujikan juga sesuai dengan karakteristik bahasa yang digunakan dalam situasi target, meskipun ragam bahasa yang digunakan hanya mewakili ragam bahasa yang digunakan dalam situasi target.

Kriteria selanjutnya adalah bahwa seksi pengujian kemahiran reseptif pada TEB BIPA 1 bermakna dalam hal variasi kebahasaan, konten, dan topik uji. Baik dalam seksi mendengar maupun membaca, konten dan topik uji yang disajikan sesuai dengan situasi yang mungkin dihadapi peuji dalam menggunakan bahasa Indonesia. Konten dan topik uji yang disajikan juga disesuaikan dengan materi yang ditentukan oleh kurikulum, sehingga sesuai dengan pembelajaran yang telah ditempuh. Dalam hal variasi kebahasaan, ragam bahasa yang digunakan adalah ragam bahasa formal dan sedikit cakapan yang disajikan. Penggunaan ragam bahasa menjadi catatan karena ragam yang diujikan sedikit berbeda dengan ragam bahasa yang digunakan dalam situasi target, khususnya dalam seksi menyimak.

Kriteria keempat adalah bahwa seksi pengujian kemahiran reseptif pada TEB BIPA 1 bermakna dalam hal mampu memprediksi kemahiran dengan kriteria eksternal, atau ekstrapolasi (Chapelle, 2021). TEB BIPA I disusun mengacu pada standar kompetensi lulusan sebagai kriteria eksternal. Kecocokan yang tinggi antara butir soal dengan kriteria tersebut dapat memprediksi capaian kemahiran peuji mengenai kompetensi apa saja yang telah dicapai. Secara rinci, kemahiran yang dicapai peuji dapat dideskripsikan melalui indikator pencapaian kompetensi yang dirancang pada setiap butir soal.

Selanjutnya, kriteria kebermaknaan yang kelima adalah bahwa seksi pengujian kemahiran reseptif pada TEB BIPA 1 konsisten dalam hal set butir/tugas dari konstruk yang berbeda. Dalam hal ini, pengembangan pengujian kemahiran reseptif pada TEB BIPA I mempunyai acuan yang terstandar dalam pengembangan paket soal yang lainnya. Acuan tersebut berupa kisi-kisi yang secara detail menentukan kompetensi, indikator, materi, rambu-rambu, dan tingkat kesukaran soal sehingga paket soal yang disusun setara satu sama lainnya. Dalam tahap pengembangan juga diselenggarakan penilaian dari pakar untuk menyetarakan tiap butir soal baik antarpaket soal maupun dengan kurikulum acuan.

Kriteria keenam dalam aspek kebermaknaan adalah bahwa seksi pengujian kemahiran reseptif pada TEB BIPA 1 konsisten dalam tugas uji yang beragam. Konsistensi ragam tugas uji dapat dilihat dari konsistensi internal atau reliabilitas uji (Brown & Lee, 2015) baik dari hasil uji coba maupun dari hasil tes yang sebenarnya. Dalam konteks pengujian kemahiran reseptif TEB BIPA I, validitas dan reliabilitas uji dari hasil uji coba terbatas menunjukkan reliabilitas yang relatif tinggi (di atas 0,76). Hal tersebut menunjukkan adanya keandalan antarkomponen dalam tes.

Kriteria terakhir dari aspek kebermaknaan adalah bahwa seksi pengujian kemahiran reseptif pada TEB BIPA 1 konsisten dengan beragam pengujian yang terlibat dalam asesmen. Dalam pengujian kemahiran reseptif, potensi perbedaan penilaian antarpengujian akan sangat kecil karena kunci jawaban telah disediakan dengan jawaban benar hanya satu pada tiap butir soal. Bobot skor yang digunakan juga menggunakan bobot skor biner dengan jawaban salah berbobot 0 dan jawaban benar berbobot 1. Pengembang juga mengantisipasi prosedur penskoran dengan menyediakan pedoman penskoran yang tercantum dalam petunjuk teknis.

Saat ini belum tersedia penelitian lanjutan mengenai proses kognitif yang disasar dalam kisi-kisi dan tingkat kesukaran butir soal yang dirancang dengan hasil yang diharapkan. Selain itu, tantangan bagi pengembang adalah dalam mewujudkan situasi target khususnya pada pengujian kemahiran reseptif. Tantangan selanjutnya adalah untuk mewujudkan kebermaknaan dalam memprediksi kemahiran dengan kriteria eksternal, perlu dicantumkan deskripsi kemahiran berdasarkan jawaban pejuang dalam laporan hasil uji untuk melihat kompetensi yang telah tercapai atau belum. Saat ini, laporan hasil uji hanya berupa predikat dan persentase skor dari tiap kemahiran. Penilaian dari pakar perlu dilaksanakan secara komprehensif untuk menyelaraskan butir soal, dan didukung dengan hasil uji coba yang memadai. Hasil uji coba perlu memadai dan memenuhi kriteria statistik untuk kalibrasi butir soal. Salah satu kriteria adalah jumlah sampel uji coba, yaitu sekitar sepuluh hingga dua puluh kali lipat dari jumlah butir soal (Andrich & Marais, 2019). Aspek keadilan juga menjadi pertimbangan untuk membobot soal secara biner untuk semua tingkat kesulitan sehingga perlu dipikirkan untuk mengadaptasi uji kemahiran bahasa adaptif dan menggunakan pembobotan soal berdasarkan tingkat kesulitan soal (Andrich & Marais, 2019),

Berdasarkan paparan tersebut, seksi pengujian kemahiran reseptif TEB BIPA I dapat dikatakan bermakna dengan memenuhi kriteria bahwa: (1) butir soal sesuai dengan acuan, (2) situasi uji setidaknya telah mewakili situasi target, (3) variasi kebahasaan, konten, dan topik uji telah disesuaikan dengan acuan dan situasi target, serta (4) hasil uji mampu memprediksi kemahiran dan selaras dengan kriteria eksternal. Seksi pengujian kemahiran reseptif TEB BIPA I juga dapat dikatakan konsisten dengan memenuhi kriteria bahwa. set uji dapat dikatakan konsisten dalam (1) konstruk yang berbeda, (2) set uji yang beragam, dan (3) pengujian/pemberi skor yang beragam. Tantangan yang dihadapi pengembang dalam menyusun tes yang lebih bermakna adalah pada kesesuaian situasi uji dengan situasi target, deskripsi kemahiran pejuang secara detail pada laporan hasil uji dengan merujuk pada kriteria eksternal/acuan, serta jumlah sampel untuk hasil uji yang memadai secara statistik. Argumentasi komprehensif diperlukan dengan memberikan penelitian tersendiri seperti analisis kecocokan proses kognitif atau tingkat kesukaran antara rancangan dengan spesifikasi dengan hasil uji, kalibrasi tes dengan menggunakan prosedur statistik yang mutakhir (misalnya menggunakan teori respons butir), serta penggunaan bobot soal yang sesuai dengan tingkat kesukaran soal.

3. Seksi pengujian kemahiran reseptif pada TEB BIPA I terbebas dari bias

Dalam memenuhi klaim mengenai ketiadaan bias, TEB BIPA 1 perlu memenuhi tiga kriteria mengenai ketiadaan bias, yaitu (1) terhindar dari bias dalam hal konten, topik, atau ragam bahasa untuk seluruh kelompok pejuang, (2) terhindar dari perbedaan kinerja karena perbedaan gender, usia, ras/etnis, atau bahasa jati, dan (3) interpretasi skor TEB BIPA 1 berdasarkan prosedur terstandar yang dapat dipertanggungjawabkan.

Kriteria pertama dari sub klaim ini adalah bahwa TEB BIPA 1 perlu terhindar dari bias dalam hal konten, topik, atau ragam bahasa untuk seluruh kelompok pejuang. Dalam pengujian

kemahiran reseptif TEB BIPA I, baik topik, konten maupun ragam bahasa yang digunakan telah disesuaikan dengan acuan yang telah ditetapkan dan disesuaikan dengan situasi target yang akan dihadapi oleh peuji dalam menggunakan bahasa Indonesia. Potensi bias masih dapat muncul khususnya dalam ragam bahasa pada situasi uji dengan ragam bahasa pada situasi target (Llosa, 2020). Hal ini sesuai dengan hasil analisis kebermaknaan juga bahwa situasi uji menggunakan ragam yang berbeda dengan situasi target.

Selanjutnya, kriteria kedua dari aspek ketiadaan bias adalah bahwa TEB BIPA 1 terhindar dari pembedaan kinerja karena perbedaan gender, usia, ras/etnis, atau bahasa jati. Dalam memenuhi kriteria ini, analisis DIF (*differentiate item functioning*) (Andrich & Marais, 2019) dilakukan untuk melihat performa butir soal terhadap kelompok peuji tertentu secara statistik. Hasil analisis pada 15 sampel uji dengan menggunakan variabel gender dan bahasa jati menunjukkan tidak adanya butir soal yang terindikasi membedakan kinerja peuji pada kedua variabel tersebut dengan probabilitas $DIF > 0,05$.

Kriteria terakhir dari subklaim ini adalah bahwa interpretasi skor TEB BIPA 1 berdasarkan prosedur terstandar yang dapat dipertanggungjawabkan. TEB BIPA I khususnya dalam seksi pengujian kemahiran reseptif menggunakan bobot skor biner 0 untuk jawaban salah dan 1 untuk jawaban benar. Skor akhir pada tiap kemahiran dipersentasekan sehingga rentang skor yang diperoleh berkisar antara 0–100. TEB BIPA termasuk penilaian acuan patokan (*criterion referenced*) yang mengacu pada kriteria standar. Interpretasi skor mencakup predikat (sempurna, sangat baik, baik, dst.) berdasarkan persentase capaian keseluruhan (Badan Pengembangan dan Pembinaan Bahasa, 2020). Kriteria yang diacu termaktub dalam SKL BIPA (Permendikbud No. 27 tahun 2017) sehingga dapat dikatakan bahwa interpretasi skor TEB BIPA melalui prosedur terstandar dapat dipertanggungjawabkan. Catatan dari interpretasi skor adalah perlu adanya deskripsi rinci mengenai kompetensi yang dicapai peuji pada laporan hasil uji.

Meskipun seksi pengujian kemahiran reseptif TEB BIPA I telah mengacu pada standar dan kurikulum acuan, potensi bias dalam konten, topik, maupun ragam bahasa *bisa saja* muncul dalam standar dan kurikulum acuan itu sendiri sehingga perlu dilakukan analisis yang komprehensif dalam mengidentifikasi potensi bias baik dalam tes maupun dalam kurikulum acuan. Sebuah studi pada bahan ajar acuan BIPA oleh Rinjaya (2020) melaporkan adanya bias gender yang cukup signifikan dalam buku ajar acuan BIPA bagi pelajar. Selain itu, analisis yang komprehensif seperti analisis DIF perlu dilakukan untuk mengidentifikasi performa butir soal terhadap kelompok peuji secara spesifik sehingga potensi bias yang dapat membedakan kelompok peuji dapat diketahui. Interpretasi skor TEB BIPA I juga mengacu pada standar yang dapat dipertanggungjawabkan yaitu kurikulum BIPA dan SKL yang termaktub dalam Permendikbud No. 21 tahun 2017. Dalam hal ini, seksi pengujian TEB BIPA I belum sepenuhnya memenuhi klaim terbebas dari bias.

4. TEB BIPA I menyediakan akses dan penyelenggaraan yang layak

Dalam memenuhi klaim mengenai aksesibilitas, TEB BIPA 1 perlu memenuhi tujuh kriteria mengenai aksesibilitas, yaitu (1) keterjangkauan harga, (2) lokasi yang mudah diakses, (3) dapat diakses oleh orang dengan kebutuhan khusus dan menyediakan akomodasi yang layak, (4) dapat diakses oleh peuji dengan bahasa jati bukan bahasa Indonesia untuk konten subjek, (5) penyelenggaraan yang seragam, (6) diselenggarakan tanpa ada kecurangan atau pelanggaran pada aspek keamanan, serta (7) pengambilan keputusan TEB BIPA 1 berdasarkan pada landasan yang kuat, termasuk landasan hukum dan etikanya.

TEB BIPA 1 pada saat ini tersedia secara gratis sesuai dengan permintaan dari lembaga penyelenggara program BIPA, sehingga TEB BIPA 1 dapat dikatakan terjangkau. Lembaga yang membutuhkan fasilitasi untuk TEB BIPA mengirimkan surat kepada pusat tes (Pustanda) untuk menyelenggarakan tes, selanjutnya tes akan diselenggarakan di lokasi yang telah disepakati baik pihak pemohon maupun pihak penyelenggara. Selain itu, mengenai lokasi TEB BIPA, TEB BIPA direncanakan akan dilaksanakan dengan tiga mekanisme, yaitu di tempat uji terpusat, tempat yang disepakati oleh pemohon dan

penyelenggara, serta secara daring. Tempat uji terpusat berada di Badan Pengembangan dan Pembinaan Bahasa termasuk Balai/Kantor Bahasa di setiap provinsi di Indonesia. Saat ini juga tengah dirancang penempatan tempat uji terpusat di seluruh Perwakilan Republik Indonesia melalui Atase Pendidikan dan Kebudayaan. Tidak menutup kemungkinan bahwa TEB BIPA akan diselenggarakan di universitas/lembaga penyelenggara program BIPA lainnya di luar negeri. Selanjutnya, untuk memenuhi kriteria aksesibilitas ketiga, TEB BIPA 1 dapat diakses oleh orang dengan kebutuhan khusus dan menyediakan akomodasi yang layak. Hal tersebut perlu diartikulasikan dalam pedoman penyelenggaraan tes dengan mencantumkan akomodasi bagi calon peuji berkebutuhan khusus.

Selain itu, TEB BIPA 1 perlu dapat diakses oleh peuji dengan bahasa jati bukan bahasa Indonesia untuk konten subjek. TEB BIPA I menyuguhkan materi/konten mengenai informasi umum seperti pengenalan, keluarga, dan deskripsi sekitar sehingga TEB BIPA dapat diakses oleh pengguna pada umumnya. Capaian kompetensi pada TEB BIPA I lebih pada ranah tujuan sintas dengan membekali penutur non jati bahasa Indonesia agar dapat berkomunikasi dengan menggunakan bahasa Indonesia ketika berkunjung ke Indonesia, baik untuk tujuan wisata, bisnis, studi, maupun tujuan lainnya. Oleh karena itu, konten-konten tersebut sebelumnya telah dikuasai oleh peuji dalam bahasa jatinya sehingga potensi bias konten dapat dihindari, khususnya dalam TEB BIPA I.

Dalam hal penyelenggaraan, TEB BIPA 1 perlu memenuhi kriteria penyelenggaraan yang seragam terlepas dari tempat penyelenggaraan tes. Keseragaman dalam penyelenggaraan tes diperlukan agar tes diselenggarakan secara setara, adil, dan tidak merugikan peuji. Ketidakeragaman penyelenggaraan dapat menguntungkan peuji di suatu tempat dan merugikan peuji di tempat lainnya. Pengembang perlu menyusun pedoman penyelenggaraan uji yang komprehensif dan mencakup segala aspek penyelenggaraan ujian dari mulai pendaftaran, pelaksanaan, hingga prosedur pengambilan laporan hasil uji. Pedoman ini juga mengatur aspek keamanan, bagaimana penyelenggara mengantisipasi potensi-potensi kecurangan yang dilakukan peuji maupun proktor uji/pengawas. TEB BIPA 1 perlu diselenggarakan tanpa ada kecurangan atau pelanggaran pada aspek keamanan. Aspek keamanan menjadi salah satu hal yang diperhatikan baik itu dalam penyelenggaraan tes secara luring berbasis kertas, luring berbasis komputer, maupun daring (Sabbah, 2017).

Kriteria terakhir dalam subklaim aksesibilitas adalah bahwa pengambilan keputusan TEB BIPA 1 berdasarkan pada landasan yang kuat, termasuk landasan hukum dan etikanya. TEB BIPA I mempunyai landasan hukum Permendikbud No. 27 tahun 2017 dengan payung besarnya Undang-undang No. 24 Tahun 2009 mengenai Bendera, Bahasa, dan Lambang Negara, serta Lagu Kebangsaan. Kriteria yang diacu TEB adalah standar kompetensi lulusan BIPA dari Permendikbud No. 27 Tahun 2017 tersebut.

TEB BIPA secara daring menimbulkan potensi ketidakadilan bagi peuji karena setiap peuji mempunyai pengalaman yang berbeda dalam mengambil tes, baik dalam hal spesifikasi komputer yang digunakan, sistem operasi, dan keandalan jaringan internet peuji dibandingkan dengan set lokasi yang terkontrol. Mekanisme pengawasan pada tes daring pun menjadi catatan dalam hal keamanan uji (Holden, Norris, & Kuhlmeier, 2021). Selain itu, sampai saat ini, aksesibilitas bagi peuji dengan kebutuhan khusus pun perlu menjadi perhatian dan dicantumkan dalam pedoman penyelenggaraan tes. Dalam aspek aksesibilitas, dapat dikatakan bahwa TEB BIPA I belum sepenuhnya menyediakan aksesibilitas dan penyelenggaraan yang layak.

SIMPULAN

Tulisan ini meninjau seksi pengujian kemahiran reseptif pada Tes Evaluasi Belajar Bahasa Indonesia bagi Penutur Asing Tingkat I (TEB BIPA I) dalam sudut pandang pendekatan berbasis etik (Kunnan, 2020). Aspek yang ditinjau adalah penerapan prinsip *fairness* sebagai salah satu landasan etik dalam pendekatan ini. Penerapan prinsip *fairness* mencakup empat aspek yaitu kesempatan belajar, kebermaknaan, ketiadaan bias, dan aksesibilitas. Klaim mengenai kesempatan belajar diartikulasikan bahwa “sebelum mengambil seksi pengujian kemahiran reseptif pada TEB BIPA I, tersedia kesempatan

belajar dan persiapan yang memadai". TEB BIPA 1 dapat dikatakan hampir memenuhi kriteria mengenai kesempatan belajar yang memadai, waktu persiapan tes yang memadai, dan latihan dengan teknologi terkini, dan pengalaman semiotika yang memadai. Hal yang menjadi catatan bahwa dalam memenuhi klaim mengenai kesempatan belajar, penyelenggara tes perlu memberikan informasi yang lengkap dan rinci kepada calon peuji, serta menyusun panduan pelaksanaan tes yang komprehensif bagi calon peuji. Selanjutnya, seksi pengujian kemahiran reseptif pada TEB BIPA I bermakna dan konsisten. Seksi pengujian TEB BIPA I dapat dikatakan bermakna dan konsisten karena telah memenuhi semua kriteria. Tantangan yang dihadapi pengembang dalam menyusun tes yang lebih bermakna adalah pada kesesuaian situasi uji dengan situasi target, deskripsi kemahiran peuji secara detail pada laporan hasil uji dengan merujuk pada kriteria eksternal/acuan, serta jumlah sampel untuk hasil uji yang memadai secara statistik. Argumentasi komprehensif diperlukan dengan memberikan penelitian tersendiri seperti analisis kecocokan proses kognitif atau tingkat kesukaran antara rancangan dengan spesifikasi dengan hasil uji, kalibrasi tes dengan menggunakan prosedur statistik yang mutakhir (misalnya menggunakan teori respons butir), serta penggunaan bobot soal yang sesuai dengan tingkat kesukaran soal. Selanjutnya, seksi pengujian kemahiran reseptif pada TEB BIPA I telah memenuhi klaim terbebas dari bias. Hal ini ditunjukkan oleh terpenuhinya kriteria pada potensi bias konten, topik, maupun ragam bahasa bagi kelompok peuji tertentu serta hasil analisis yang komprehensif pada potensi bias butir soal yang membedakan kelompok peuji tertentu. Kriteria mengenai acuan interpretasi skor juga dapat dipenuhi. Terakhir, TEB BIPA I belum sepenuhnya menyediakan aksesibilitas dan penyelenggaraan yang layak bagi peuji. TEB BIPA secara daring menimbulkan potensi *ketidakadilan* bagi peuji karena setiap peuji mempunyai pengalaman yang berbeda dalam mengambil tes, serta aspek keamanan yang perlu menjadi perhatian khusus. Selain itu, perlu diartikulasikannya dukungan aksesibilitas bagi peuji dengan kebutuhan khusus dalam pedoman penyelenggaraan tes. Dalam hal ini, TEB BIPA I belum memenuhi semua aspek prinsip *fairness* dalam tinjauan berbasis etik, khususnya mengenai penyediaan informasi mengenai tes, petunjuk pelaksanaan tes bagi calon peuji, penelitian lanjutan mengenai kalibrasi tes, pengantisipasi potensi bias dan tantangan yang berhubungan dengan administrasi penyelenggaraan tes, seperti keamanan tes dan akomodasi peuji dengan kebutuhan khusus.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Lembaga Pengelola Dana Pendidikan Kementerian Keuangan sebagai pemberi dana hibah secara penuh dalam penelitian ini serta Pusat Penguatan dan Pemberdayaan Bahasa, Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.

DAFTAR PUSTAKA

- Andrich, D., & Marais, I. (2019). *A course in rasch measurement theory: Measuring in the educational, social and health sciences*. Singapore: Springer Singapore. <https://doi.org/10.1007/978-981-13-7496-8>
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford ; New York: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford ; New York: Oxford University Press.
- Badan Pengembangan dan Pembinaan Bahasa. (2020). *Petunjuk teknis pengembangan tes evaluasi belajar bahasa Indonesia bagi penutur asing (TEB BIPA)*. Jakarta.
- Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (Third edition). Hoboken, NJ: Pearson Education.
- Brown, H. D., & Lee, H. (2015). *Teaching by principles: An interactive approach to language pedagogy* (Fourth edition). White Plains, NY: Pearson Education.
- Carr, N. T. (2011). *Designing and analyzing language tests*. Oxford ; New York: Oxford

- University Press.
- Chapelle, C. A. (2021). *Argument-based validation in testing and assessment*. Thousand Oaks: SAGE Publications, Inc.
- Cheng, L., & Fox, J. (2017). *Assessment in the language classroom: Teachers supporting student learning*. London: Macmillan Education, Palgrave.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (Fifth edition). Los Angeles: SAGE.
- Dahler-Larsen, P. (2018). Qualitative evaluation: Methods, ethics, and politics with stakeholders. Dalam N. K. Denzin & Y. Lincoln (Ed.), *The SAGE handbook of qualitative research* (5th ed., hlm. 867–886). London; Los Angeles.: SAGE Publications.
- Davies, A. (1997). Introduction: The limits of ethics in language testing. *Language Testing*, 14(3), 235–241. <https://doi.org/10.1177/026553229701400301>
- Deygers, B., Van den Branden, K., & Van Gorp, K. (2018). University entrance language tests: A matter of justice. *Language Testing*, 35(4), 449–476. <https://doi.org/10.1177/0265532217706196>
- Elder, C. (2018). Test review: Certifying French competency: The DELF tout public (B2). *Language Testing*, 35(4), 615–623. <https://doi.org/10.1177/0265532218781627>
- Flick, U. (Ed.). (2014). *The SAGE handbook of qualitative data analysis*. Los Angeles: SAGE.
- Fulcher, G. (2010). *Practical language testing*. London: Hodder Education.
- Hoang, G. T. L., & Kunnan, A. J. (2016). Automated essay evaluation for English language learners: A case study of MY Access. *Language Assessment Quarterly*, 13(4), 359–376. <https://doi.org/10.1080/15434303.2016.1230121>
- Holden, O. L., Norris, M. E., & Kuhlmeier, V. A. (2021). Academic integrity in online assessment: A research review. *Frontiers in Education*, 6, 639814. <https://doi.org/10.3389/educ.2021.639814>
- Hughes, A. (2003). *Testing for language teachers* (2nd ed). Cambridge; New York: Cambridge University Press.
- Kane, M. T. (2004). Certification testing as an illustration of argument-based validation. *Measurement: Interdisciplinary Research & Perspective*, 2(3), 135–170. https://doi.org/10.1207/s15366359mea0203_1
- Kunnan, A. J. (1997). Connecting fairness and validation. Dalam A. Huhta, V. Kohonen, L. Kurki-Suonio, & S. Luoma (Ed.), *Current developments and alternatives in language assessment: Proceedings of LTRC 1996* (hlm. 85–109). Jyväskylä: University of Jyväskylä.
- Kunnan, A. J. (2004). Test fairness. Dalam M. Milanovic & C. Weir (Ed.), *European language testing in a global context* (hlm. 27–48). Cambridge, UK: Cambridge University Press.
- Kunnan, A. J. (2018). *Evaluating language assessments* (First edition). New York, NY: Routledge.
- Kunnan, A. J. (2020). A case for an ethics-based approach to evaluate language assessments. Dalam G. J. Ockey & B. A. Green (Ed.), *Another generation of fundamental considerations in language assessment* (hlm. 77–93). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-15-8952-2_6
- Liu, S., & Kunnan, A. J. (2015). Investigating the application of automated writing evaluation to Chinese undergraduate English majors: A case study of WriteToLearn. *CALICO Journal*, 33(1), 71–91. <https://doi.org/10.1558/cj.v33i1.26380>
- Llosa, L. (2020). Revisiting the role of content in language assessment constructs. Dalam G. J. Ockey & B. A. Green (Ed.), *Another Generation of Fundamental Considerations in Language Assessment* (hlm. 29–42). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-15-8952-2_3
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (Third edition). Thousand Oaks, California: SAGE Publications, Inc.
- Moghadam, M., & Nasirzadeh, F. (2020). The application of Kunnan's test fairness framework (TFF) on a reading comprehension test. *Language Testing in Asia*, 10(1), 7.

- <https://doi.org/10.1186/s40468-020-00105-2>
- Norris, J., & Drackert, A. (2018). Test review: TestDaF. *Language Testing*, 35(1), 149–157. <https://doi.org/10.1177/0265532217715848>
- Peng, Y., Yan, W., & Cheng, L. (2021). Hanyu Shuiping Kaoshi (HSK): A multi-level, multi-purpose proficiency test. *Language Testing*, 38(2), 326–337. <https://doi.org/10.1177/0265532220957298>
- Rinjaya, D. (2020). Representasi gender dalam buku BIPA 7 seri pelajar “Sahabatku Indonesia.” *Jurnal Bahasa Indonesia bagi Penutur Asing (JBIPA)*, 2(2), 100. <https://doi.org/10.26499/jbipa.v2i2.2958>
- Sabbah, Y. W. (2017). Security of online examinations. Dalam I. Palomares Carrascosa, H. K. Kalutarage, & Y. Huang (Ed.), *Data Analytics and Decision Support for Cybersecurity* (hlm. 157–200). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-59439-2_6
- Toulmin, S. (2003). *The uses of argument* (Updated ed). Cambridge, U.K.; New York: Cambridge University Press.
- Wagner, E. (2020). Duolingo Englishtest, revised version July 2019. *Language Assessment Quarterly*, 17(3), 300–315. <https://doi.org/10.1080/15434303.2020.1771343>
- Zhao, C. G., & Liu, C. J. (2019). An evidence-based review of Celpe-Bras: The exam for certification of proficiency in Portuguese as a foreign language. *Language Testing*, 36(4), 617–627. <https://doi.org/10.1177/0265532219849000>.