

PENGEMBANGAN TES *READING FOR ACADEMIC PURPOSES* UNTUK PROGRAM EAP DI IAIN SURAKARTA

Imroatus Solikhah
Program Studi Pendidikan Bahasa Inggris
IAIN Surakarta
Email: ratu.shyma@yahoo.com

Abstract: The objective of this study is to devise standardized test for Reading for Academic Purposes (WAP). The test was focused on WAP as the main competency of reading for EAP and elaborated sub-competencies into literal, inferential, and critical reading. The items of the test were developed from 9 passages before trials to explore critical reading as main academic literacy of EAP program. A quantitative approach was used in this study emphasizing statistical analysis for manual and SPSS program. Subjects of this study for three time trials were 30 students of EAP class at IAIN Surakarta. Each trial involved 10 students. The study revealed that of 40 items of reading test, 24 items derived from 6 passages performed accepted validation showing five criteria: difficulty indexes, degrees of difficulty, distractor effectiveness, validity, and reliability.

Key-words: reading for academic purposes, validity, reliability, test.

Tes kompetensi bahasa Inggris untuk program *English for Academic Purposes* (EAP) belum dirumuskan oleh para ahli pembelajaran bahasa Inggris secara khusus. Umumnya, tes kompetensi untuk EAP menggunakan tes terstandar yang sudah tersedia misalnya soal-soal TOEFL (*Test of English as a Foreign Language*) atau soal IELTS (*International English Language Testing Service*). Hal ini disebabkan kedua jenis tes tersebut sudah tersedia dalam jumlah dan variasi yang cukup banyak di pasaran sehingga pengguna tinggal melaksanakan saja. Selain itu, hasilnya bisa diandalkan untuk mengetahui kompetensi peserta tes.

Model penggunaan tes terstandar seperti itu memiliki keunggulan dilihat dari kepraktisan dan kegunaannya. Namun, penggunaan model tes tersebut juga memiliki kelemahan. *Pertama*, hasil tes belum tentu sesuai dengan bahan ajar yang digunakan. *Kedua*, dilihat dari segi pengembangan bahan ajar tes yang digunakan menyebabkan dosen akan menggunakan bahan ajar menyesuaikan dengan isi tes yang akan digunakan sebagai alat ukur. *Ketiga*, dilihat dari kreatifitas agar dosen mengembangkan alat ukur sendiri, penggunaan tes terstandar tidak mendorong dosen untuk melakukan sendiri alat ukur yang sesuai dengan tujuan pembelajarannya dan tidak mendorong dosen untuk meningkatkan kompetensi mengembangkan alat ukur hasil pembelajaran.

Mata kuliah EAP, sejauh ini dianggap mata kuliah kelas kedua dan kurang mendapat perhatian dalam kurikulum pendidikan bahasa termasuk di IAIN Surakarta. Mata kuliah EAP disamakan dengan mata kuliah ESP (*English for Specific Purposes*) dan pelaksanaannya dalam perkuliahan juga belum dirumuskan secara mantap.

Kajian literatur terbaru untuk EAP menunjukkan bahwa di hampir seluruh belahan dunia, perguruan tinggi telah mengembangkan EAP secara cermat. Di negara penutur bahasa Inggris, EAP telah dikemas sebagai kompetensi utama sebelum mahasiswa baik yang berasal dari negara penutur bahasa Inggris (*English speaking countries*) maupun negara pengguna bahasa Inggris sebagai bahasa kedua (*EFL, English as a Foreign Language*) mengikuti perkuliahan. Di Indonesia, EAP dilaksanakan belum standar di setiap perguruan tinggi (Solikhah, 2013).

Terdapat dua kondisi yang menunjukkan bagaimana bahasa Inggris diselenggarakan di universitas di Indonesia. *Pertama*, untuk mahasiswa jurusan non-bahasa Inggris, mata kuliah bahasa Inggris diselenggarakan dalam bentuk MKDU (Mata Kuliah Dasar Umum) yang bersifat wajib. MKDU ini diselenggarakan dalam 2-4- sks dan materi diarahkan pada penguatan isi (*content knowledge*) sesuai jurusan dan dikategorikan sebagai ESP. Misalnya, mahasiswa jurusan Biologi diajar MKDU Bahasa Inggris dengan muatan biologi. Mata kuliah ini diajar oleh dosen bahasa Inggris. Untuk contoh kasus di Prodi Biologi, Program studi Biologi menugaskan dosen Program studi Sastra Inggris atau Pendidikan bahasa Inggris untuk mengajar di kelas program studi biologi. *Kedua*, bahasa Inggris untuk jurusan Pendidikan Bahasa Inggris diajarkan dalam mata kuliah ESP atau EAP. ESP atau EAP berisi bagaimana mahasiswa merancang bahasa Inggris untuk ESP atau EAP jika sudah lulus nanti. Berbeda dengan MKDU Bahasa Inggris, mata kuliah ESP atau EAP untuk mahasiswa jurusan bahasa Inggris menekankan pada rancangan silabus, design kurikulum, bahan ajar ESP/EAP dan evaluasi EAP/ESP.

Dalam kedua kasus tersebut, ESP ialah yang dalam kajian pustaka dan berbagai penelitian lebih banyak mendapat perhatian. Bahan ajar ESP/EAP banyak disusun oleh para dosen di setiap universitas baik secara tim maupun individu, tetapi perangkat tes atau alat ukur terstandar belum ditemukan dalam kajian literatur. Penelitian ini mencoba merumuskan tes kompetensi EAP untuk mahasiswa jurusan non-bahasa Inggris. Sebagai tes yang dikembangkan sendiri yang divalidasi menggunakan acuan tes terstandar, perangkat tes ini memiliki landasan yang mantap.

Secara teoritis, tes yang dikembangkan ini mengacu pada prosedur pengembangan alat ukur yang terstandar. Suatu tes dikatakan terstandar apabila dikembangkan dan divalidasi mengikuti prosedur yang sudah ditetapkan. Prosedur pengembangannya di antaranya ialah perangkat tes tersebut diuji coba butir-butirnya dan setiap butir divalidasi melalui langkah-langkah tingkat kesulitan, daya pembeda, efektifitas pengecoh untuk soal pilihan ganda, validitas dan reliabilitas (Budiharo, 2004; Djiwandono, 2010).

Secara praktis, pengembangan tes termasuk keterampilan di bidang pengukuran dan evaluasi pembelajaran yang memerlukan ketekunan. *Pertama*, pengembangan tes yang melibatkan analisis validitas item tidak mutlak diperlukan untuk pengukuran hasil belajar harian (formatif tes). Dalam proses pembelajaran guru dianggap sudah menguasai peta pemahaman pembelajar sehingga bentuk tes yang didesain sebagai *classroom test* dianggap sudah memadai (Brown, 1997; Djiwandono, 1997). *Kedua*, pembuatan tes yang melibatkan uji-coba untuk analisis item dan validitas butir soal memerlukan waktu yang cukup panjang dan keterampilan statistik yang relatif rumit. Karena itu, tidak semua pengajar memiliki kesempatan melakukan uji coba butir soal

dan bahkan sebagian lainnya tidak telaten karena kurang menguasai teknik analisisnya (Budiharso, 2005).

Pertimbangan lain yang dijadikan dasar pemikiran penelitian ini ialah kompetensi dalam EAP. Dalam tinjauan akademis, EAP diarahkan untuk mencapai tingkat kompetensi bahasa Inggris yang disebut literasi akademis, yaitu kompetensi yang membantu pembelajar mencapai tingkat penguasaan bahasa Inggris dalam lingkungan akademis agar mahasiswa berhasil menyelesaikan perkuliahan. *Intersegmental California Academic Senate (ICAS)*, lembaga yang dibentuk pemerintah untuk menetapkan standar literasi akademik siswa dan mahasiswa di California, Amerika Serikat, membagi literasi akademik ke dalam tiga kompetensi utama, yaitu: membaca, menulis, dan berpikir kritis. Sub-kompetensi yang diwajibkan untuk mencapai kompetensi utama tersebut meliputi: membaca kritis, menulis akademik, berbicara, penguasaan kosa kata, dan penguasaan gramatika standar (ICAS, 2012:3).

Pengelompokan kompetensi yang dicakup dalam literasi akademik tersebut, dalam praktik EAP dikelompokkan ke dalam kompetensi membaca untuk tujuan akademis (*Reading for Academic Purposes*, RAP) dan menulis untuk tujuan akademis (*Writing for Academic Purposes*, WAP). Untuk itu, bahan ajar EAP menekankan pada penguasaan membaca kritis. Membaca kritis, dibagi ke dalam tiga sub-kompetensi, yaitu: *literal comprehension*, *inferential comprehension*, dan *critical comprehension*. Materi *writing* dikembangkan tersendiri untuk menguatkan kompetensi menulis akademis berupa: esai, ringkasan, laporan, artikel jurnal, makalah, tesis, dan disertasi sesuai dengan tujuan akademisnya pembelajar. Mata kuliah *critical thinking* dikembangkan dalam bentuk *speaking for academic purposes*, seperti presentasi, diskusi, seminar, rapat, dan memberikan penyajian suatu materi (ICAS, 2012; Yurekli, 2012; Kim, 2013).

Penelitian ini mengkaji khusus untuk pengembangan tes kompetensi membaca untuk EAP. Proses pengembangan instrumen yang di dalamnya mengharuskan peneliti untuk melakukan uji coba. Langkah-langkah pengembangan instrumen penelitian tersebut sudah baku dan secara ketat harus dilakukan peneliti. Tujuan pelaksanaan uji coba tersebut ialah terutama variabel yang diukur melalui indikator yang harus diidentifikasi dan dijelaskan secara tepat oleh peneliti. Melalui pengembangan instrumen, peneliti juga dituntut merumuskan kerangka pikir penelitian, definisi konsep dan definisi operasional yang akan menuntun peneliti dalam mengambil data. Karena itu, esensi pelaksanaan uji-coba butir soal sangat penting dalam penelitian dan peneliti dituntut untuk melakukan rangkaian uji coba instrumen melalui tahapan-tahapan yang terstandar.

Rangkaian uji coba yang dilakukan meliputi tahap pengembangan butir, uji-coba pertama untuk melihat tingkat kesulitan, dan uji-coba kedua untuk melihat validitas dan reliabilitas butir. Membaca pemahaman dipilih sebagai fokus karena pertimbangan praktis saja. Selain teks mudah dikembangkan dalam bentuk butir soal, peneliti memiliki pengalaman empiris dalam hal mengajar membaca pemahaman. Kedua alasan tersebut dianggap memudahkan peneliti untuk memudahkan menyusun butir soal dan menganalisis hasil uji coba.

Membaca hakikatnya ialah proses membangun pemahaman wacana tulis. Proses ini terjadi dengan menjodohkan atau menghubungkan skemata pengetahuan dan

pengalaman yang telah dimiliki sebelumnya sehingga membentuk pemahaman terhadap wacana yang dibaca (Syafiie, 1999:35). Pemahaman dalam membaca dikelompokkan ke dalam kelompok membaca rendah dan kelompok membaca tinggi. Kelompok membaca rendah disebut juga membaca literal. Kelompok membaca tinggi disebut juga membaca kritis, dan terdiri dari tiga: pemahaman interpretif, pemahaman kritis, dan pemahaman kreatif (Burns, Roe, dan Ross, 1996). Tindak membaca terdiri dari dua bagian: proses membaca dan hasil membaca. Sebagai proses, membaca melibatkan juga aktivitas mental dan fisik. Sedangkan hasil membaca mengacu pada komunikasi pikiran atau perasaan penulis kepada pembaca.

Pemahaman yang paling rendah ialah membaca pemahaman literal. Pemahaman literal mengacu pada apa yang disebutkan dalam teks bacaan. Tujuan pemahaman literal ialah untuk memperoleh detail isi bacaan secara efektif. Pemahaman interpretif dihasilkan melalui proses berpikir yang lebih tinggi, yaitu interpretasi, analisis, dan sintesis suatu bacaan. Pemahaman interpretif menuntut pembaca membuat interpretasi isi wacana walaupun informasi yang diperlukan tidak terdapat dalam teks. Pemahaman dibentuk dengan cara membuat analisis atau inferensi. Dalam pemahaman jenis ini prediksi isi bacaan ialah keterampilan yang penting, sebab dengan memprediksi seseorang dapat membuat hipotesis dilanjutkan dengan penerimaan, pengakuan, atau penolakan terhadap isi hipotesis tersebut.

Pemahaman kritis bertujuan untuk memberi penilaian terhadap isi teks yang dibaca. Penilaian dilakukan dengan cara membandingkan dengan pengetahuan yang sudah dimiliki atau sumber lain. Dalam membaca jenis ini, pembaca bisa memberi komentar, menemukan kekurangan, atau menjelaskan keunggulan isi teks.

Tingkat pemahaman kreatif menuntut pembaca untuk menemukan ide dalam teks yang dianggap kurang benar, kurang akurat, atau bertentangan dengan sumber lain. Jika dalam membaca kritis, pembaca cukup memberi komentar atau pendapat lain, dalam membaca kreatif, pembaca bisa membuat telaah, summary, atau kompilasi hasil bacaan.

Pertanyaan dalam membaca pemahaman yang digunakan dalam laporan ini mengacu pada rumusan yang dikembangkan oleh Shape (2000:316) dalam Baron's TOEFL. Shape (2000) mengidentifikasi delapan jenis pertanyaan yang lazim digunakan dalam tes membaca. Ke-8 jenis tersebut ialah: preview (*previewing*), pertanyaan untuk mengetahui ide pokok (*reading for main ideas*), pertanyaan kosa kata dalam konteks (*using context for vocabulary*), skaning untuk mencari detail (*scanning for details*), membuat inferensi (*making inferences*), mengidentifikasi perkecualian (*identifying exceptions*), menemukan referens (*locating references*), dan pertanyaan yang merujuk kembali pada teks (*referring to the passage*).

Pertanyaan preview bertujuan untuk mengetahui ide umum suatu teks. Pertanyaan ini bertujuan untuk mengetahui ide umum yang tertuang dalam teks; untuk mengetahui kesan mengenai isi teks. Preview biasanya dikerjakan dengan cara membaca kalimat awal dan kalimat akhir pada setiap paragraf.

Membaca untuk mengetahui ide utama bertujuan untuk mengetahui pandangan atau maksud penulis. Isi pesan bisa ditangkap melalui pernyataan *thesis* dalam teks. Secara spesifik, ide utama bisa diketahui dari topik apa yang dibicarakan oleh penulis dalam teks. Jenis pertanyaan untuk mengetahui ide utama bisa berupa: (1) Apa ide

utama teks?, (2) Apa subjek yang ditulis?, (3) Apa topik bacaan?, dan (4) Judul apa yang paling cocok?

Dalam pertanyaan makna kosa kata yang mengacu ada konteks, pertanyaan bisa diketahui jawabannya dengan melihat pada kata, frase, atau kalimat yang ada di sekitar kata yang ditanyakan maknanya. Itulah yang disebut konteks, yang berupa kata atau kalimat dalam bacaan. Konteks berfungsi untuk menebak makna. Umumnya, jika pembaca mengetahui makna umum suatu kalimat, dia juga akan tahu makna umum dalam kalimat. Jadi, menebak makna melalui konteks bisa digunakan untuk menebak makna kosa kata tertentu yang ditanyakan dalam soal.

Dalam scanning, pertanyaan diarahkan untuk mengetahui informasi detail tertentu. Untuk menjawab pertanyaan jenis ini, ada tiga langkah yang diperlukan. Pertama, pembaca harus menemukan *content words* berupa kata benda, kata kerja, kata sifat, dan kata keterangan yang dianggap penting. Kedua, pembaca harus mencari kata yang memiliki sinonim atau antonim dari kata yang dicari. Dengan cara membaca cepat, pembaca akan menemukan tempat kata yang ditanyakan dalam teks. Selanjutnya, pembaca harus mengkaji kalimat yang ditanyakan sesuai dengan bunyi pertanyaannya.

Pertanyaan inferensi ialah pertanyaan yang keterangannya tidak dinyatakan langsung dalam teks. Jadi pembaca harus mengaitkan dengan keterangan dalam teks atau informasi di luar teks. Biasanya yang terdapat dalam teks ialah fakta atau pernyataan tertentu yang bisa dikaitkan satu sama lainnya. Jadi pertanyaan inferensi dibuat berdasarkan penarikan simpulan logis menggunakan fakta dalam teks atau pendapat penulis.

Pertanyaan yang mengacu pada pengecualian berisi beberapa pernyataan yang dianggap semua salah atau semua benar; dan dari pernyataan tersebut terdapat satu yang merupakan pengecualian yang merupakan jawaban yang diminta. Pertanyaan jenis ini bisa dijawab menggunakan teknik skaning.

Mengidentifikasi referensi dimaksudkan untuk mengetahui pelaku atau kata ganti yang sesuai dengan konteks. Umumnya, pertanyaan dimulai dengan kata ganti: *it*, *he*, *she*, dsb, dan pembaca diminta untuk menetapkan kata ganti yang paling cocok dengan konteks. Teknik menjawab pertanyaan jenis ini meliputi menemukan lebih dulu kata ganti yang ditanyakan dan temukan kata ganti dalam ada dalam pertanyaan dan pilihan jawaban. Makna kalimat tidak akan berubah kalau pembaca menemukan kata ganti yang tepat.

Jenis pertanyaan terakhir ialah menetapkan isi teks. Pertanyaan jenis ini ditandai dengan cara menanyakan makna baris tertentu dalam teks. Biasanya, pertanyaannya digaris bawahi. Teknik yang bisa digunakan untuk menjawab pertanyaan ialah pembaca harus mengidentifikasi lebih dulu bunyi pertanyaannya, lalu temukan makna kata atau kalimat yang digarisbawahi dalam teks.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan analisis statistik sebagai alat analisis utama. Prosedur penelitian dilakukan mulai dari pengembangan butir, uji coba lapangan, dan analisis. Pengembangan butir dimulai dari sintesis teori pendukung, penyusunan definisi konsep, perumusan definisi operasional, pembuatan kisi-kisi, dan pengembangan butir soal. Setelah itu, butir soal diujicobakan beberapa

kali. Ujicoba pertama digunakan untuk mengidentifikasi tingkat kesulitan butir (*item degrees of difficulty*), daya pembeda (*discrimination*), efektivitas distraktor (untuk butir soal dengan bentuk pilihan ganda). Butir soal yang dianggap sesuai dengan kriteria tingkat kesulitan yang diperoleh setelah disusun ulang, diujicobakan lagi untuk mengetahui tingkat validitas dan reliabilitas. Kerangka pikir ini mudah diikuti terutama yang menggunakan analisis secara manual. Walaupun rumit, analisis ini memiliki keunggulan karena pelacakan butir bisa lebih teliti, mulai dari tingkat kesulitan, daya pembeda, dan distraktor. Penelitian ini melakukan analisis butir dua kali: manual dan program SPSS Release 11. Kedua model analisis ini bisa memberikan informasi lengkap yang meliputi: tingkat kesulitan butir, daya pembeda, efektivitas distraktor, validitas, dan reliabilitas.

HASIL

Proses Ujicoba

Ujicoba butir soal ini dilakukan sebanyak tiga kali. Setiap kelompok ujicoba digunakan 10 responden. MAN 1 Blitar dipilih karena pertimbangan kepraktisan dan akses. Ujicoba pertama, dimaksudkan untuk menentukan kualitas butir yang meliputi: tingkat kesulitan butir, daya pembeda, dan afektivitas distraktor. Peserta yang digunakan sebagai responden ialah kelompok A. Dari hasil ujicoba pertama, diperoleh butir soal sebanyak 25 butir; yang berarti 5 butir dari sebelumnya berjumlah 30 dibuang. Kelima butir yang dibuang tersebut ialah seluruh pertanyaan pada teks pertama. Setelah hasil analisis kualitas butir diketahui, butir soal direkonstruksi lagi dan diperoleh 25 butir yang dikembangkan dari 6 teks. Butir soal ini dikemas lagi menjadi seperangkat tes dan diujicobakan lagi kepada 10 orang kelompok B. Tujuan ujicoba yang kedua ini ialah untuk mengetahui validitas dan reliabilitas ke-25 butir soal tersebut (Brown, 1995; Djiwandono, 1995). Uji validitas dilakukan menggunakan teknik kriteria; nilai ujicoba dibandingkan dengan nilai guru yang diperoleh dari tes membaca pemahaman. Sedangkan uji reliabilitas dilakukan menggunakan teknik test-retest. Untuk memperoleh skor uji reliabilitas, tes yang sama diberikan kepada peserta yang sama dalam waktu yang hampir bersamaan (Djiwandono, 1995). Dengan demikian, siswa kelas III-B dites sebanyak 2 kali dalam waktu berselang 30 menit dari tes pertama dan tes kedua.

Karena analisis ini menggunakan dua teknik, manual dan SPSS Release 11, peneliti secara sengaja mengadakan tes yang keempat kalinya. Tes ini diberikan pada kelompok C. Butir tes yang digunakan ialah butir soal yang sudah dianalisis kualitas butirnya sebanyak 24 item. Kelas A, B, dan C sengaja dites terpisah untuk menghindari kebocoran pertanyaan dan menjaga validitas eksternal. Selain itu, analisis yang digunakan ialah manual dan SPSS, karena itu, diperlukan jenis data yang berbeda.

Teknik Manual

Teknik manual digunakan untuk mengetahui lima hal: tingkat kesulitan butir item, daya pembeda item, efektivitas distraktor, validitas, dan reliabilitas. Tingkat kesulitan butir item menunjukkan taraf kesulitan butir dalam tiga katagori: sedang, sulit, dan sangat sulit. Menggunakan kriteria dari Djiwandono (1995) dan Brown (1995), kriteria tingkat kesulitan yang digunakan ialah 20% sd 80%. Semakin tinggi

persentasenya menunjukkan butir mudah dan semakin rendah mendekati 20% menunjukkan butir soal sulit. Kriteria tingkat kesulitan yang diusulkan Djiwandono (1995) dan Brown (1995) ialah:

81-100%	Sangat mudah
70-80	Mudah
41-69	Sedang
20-40	Sulit
10-20	Sangat sulit

Daya pembeda menunjukkan kemampuan butir soal dalam membedakan peserta tes yang mengerti dan yang tidak mengerti. Mengikuti Dhiwandono (1995) dan Brown (1995), kriteria diskriminasi yang digunakan ialah 20% sd 100%. Kebalikan dari tingkat kesulitan, semakin mendekati 100%, daya pembeda butir semakin bagus. Brown (1995) dan Djiwandono (1995) mengusulkan kriteria diskriminasi berikut ini.

100%	Diskriminasi sempurna
50%	Ideal
20%	Sedang
0%	Tidak ada diskriminasi
(-) Negatif	Diskriminasi terbalik (kelompok rendah menjawab benar dan kelompok tinggi justru menjawab salah)

Analisis validitas yang digunakan ialah validitas kriteria. Validitas maksudnya kesesuaian atau kecocokan antara nilai tes dan kemampuan peserta tes (Djiwandono, 1995). Dalam laporan ini, analisis validitas kriteria memerlukan dua set skor yang diperoleh dari skor tes ujicoba dan skor pembandingan dari tes lain yang menguji hal yang sama. Dalam laporan ini, skor pembandingan diperoleh dari nilai kelompok B yang mengikuti ujicoba kedua. Skor diberikan oleh guru yang juga diperoleh dari nilai membaca pemahaman. Analisis digunakan rumus Product Moment.

Analisis reliabilitas yang digunakan dalam laporan ini ialah analisis menggunakan teknik test-retest. Reliabilitas maksudnya keajegan atau konsistensi dari skor pada tes pertama dibandingkan skor pada tes kedua. Skor tes pertama diperoleh dari tes pada ujicoba kedua yang diberikan pada kelompok B. Adapun skor kedua diperoleh juga dari kelompok B estela mereka diberi waktu 25 menit untuk istirahat dari tes pertama, untuk mengerjakan tes yang sama yang kedua kalinya. Sama dengan analisis validitas, analisis reliabilitas juga menggunakan humus product Moment. Kriteria penentuan validitas dan reliabilitas dinyatakan dalam bentuk persen untuk memudahkan interpretasi berapa persen tingkat akurasi. Kriteria yang dikemukakan Djiwandono (1995) dan Brown (1995) disajikan berikut ini.

80%-100%	Sangat tinggi
70%-79%	Tinggi
60%-69%	Sedang
40%-59%	Kurang
10%-39%	Sangat kurang

Analisis Menggunakan SPSS Release 11

Taknik analisis menggunakan program SPSS Release 11 digunakan sebagai pembandingan. Analisis ini bertujuan untuk mengetahui validitas butir dan reliabilitas total seluruh item. Program SPSS Release 11 dipilih karena program ini termasuk yang mutakhir dan peneliti menguasai program ini.

Analisis menggunakan SPSS Release 11 memerlukan cukup satu set skor. Skor tersebut diambil dari tes yang sudah diseleksi tingkat kesulitannya sebanyak 24 butir. Untuk membedakan dengan hasil analisis manual, peneliti sengaja menggunakan data yang berbeda yang diperoleh dari ujicoba pada kelompok C. Hasil analisis SPSS memberikan informasi sejauh mana hasil uji korelasi setiap butir lebih besar atau lebih kecil dari kriteria. Jika hasil uji statistik suatu butir lebih kecil dari kriteria, butir tersebut dianggap tidak valid dan tidak bisa digunakan. Sebaliknya, hasil uji statistik butir soal yang lebih tinggi dari kriteria, dianggap valid dan direkomendasikan untuk digunakan.

Hasil Uji Coba Menyeluruh

Model tes membaca dan kosa kata ini mengadopsi model tes TOEFL atau IELTS. Secara operasional yang dimaksud tes membaca pemahaman di sini ialah tes untuk mengetahui ketarampilan memahami isi teks melalui teknik pemahaman literal, inferensial, dan pemahaman kritis. Yang dimaksud tes kosa kata ialah tes untuk menguji pemahaman makna kosa kata akademik dan kosa kata umum 1.000-2.000 dan kosa kata akademik.

Awalnya tes berjumlah 40 butir, setelah ujicoba terseleksi 24 butir. Tes ini diujicobakan kepada 10 mahasiswa EAP di IAIN Surakarta. Ujicoba bertujuan untuk mengkaji tingkat kesulitan butir (p), daya beda atau diskriminasi (D), keefektifan pengecoh, validitas, dan reliabilitas.

Karakteristik Tes

Tes Kompetensi Bahasa Inggris setelah ujicoba berjumlah 24, berbentuk pilihan ganda dengan 4 pengecoh. Materi tes meliputi membaca pemahaman dan kosa kata. Sebelum uji-coba perangkat tes direview oleh pakar tes bahasa untuk melihat kesesuaian jumlah butir dengan materi EAP, keragaman butir, pilihan, dan proporsi materi membaca dan kosa kata. Teknik pengetikan, penyajian urutan butir dan akurasi petunjuk mengerjakan soal. Secara khusus review diarahkan untuk melihat: *stem* atau *lead* (pertanyaan), pengecoh, homogenitas dan kunci jawaban soal.

Tingkat Kesulitan dan Daya Beda Butir Soal

Cakupan analisis meliputi: (1) tingkat kesulitan, daya beda, dan pengecoh, (2) validitas butir soal, dan (3) reliabilitas.

Tabel 1. Profil Soal Setelah Kalibrasi

No	Jenis	Nomor Item	Jumlah	Persentase
1	Reading: Teks 1	1-5	5	11.1
2	Reading: Teks 2	6-9	4	8.9
3	Reading: Teks 3	10-13	4	8.9
4	Reading: Teks 4	14-17	4	8.9
5	Reading: Teks 5	18-21	4	8.9
6	Reading: Teks 6	22-24	3	6.7
			45	100

Keefektifan Pengecoh

Keefektifan pengecoh ialah sejauh mana setiap pengecoh berfungsi efektif mengelabui peserta tes. Keefektifan ditunjukkan dengan setiap pengecoh harus dipilih oleh minimal satu peserta tes. Hasil analisis menunjukkan secara umum pengecoh berfungsi baik. Revisi dilakukan untuk pengecoh D pada butir 54 karena pengecoh tidak ada yang memilih dan kunci jawaban untuk butir nomor 20 dipilih lebih sedikit dari pada pengecoh.

Validitas

Validitas ialah kesesuaian antara butir soal dengan skor yang diperoleh peserta tes. Uji validitas penelitian ini ialah korelasi biserial dan dilakukan bersama-sama dengan analisis tingkat kesulitan, daya beda, dan keefektifan pengecoh. Dari 60 butir soal yang dibuat, terdapat 15 butir yang dinyatakan tidak valid.

Reliabilitas

Reliabilitas ialah keajegan atau konsistensi skor yang diperoleh dari setiap kali tes digunakan. Reliabilitas dalam penelitian ini dianalisis menggunakan Rumus KR-20. Hasil analisis menunjukkan $r = 0.829$ yang berarti butir soal penelitian ini memiliki konsistensi sampai 82.9%.

BAHASAN

Hasil analisis ujicoba tes membaca pemahaman literal ini menunjukkan bahwa dari 40 butir soal yang dianalisis, terdapat 24 butir yang dianggap valid dan bisa digunakan dalam tes yang sebenarnya. Proses pembuktian melalui ujicoba ini menggunakan perpaduan teknik manual dan teknik program SPSS.

Untuk memperoleh hasil ujicoba yang baik, diperlukan beberapa kali ujicoba. Laporan ini menggunakan empat kali tes yang diberikan terhadap siswa yang berbeda. Ujicoba yang pertama diberikan pada 10 orang yang diseleksi secara random. Tujuan pemberian ujicoba ini ialah untuk mengetahui kualitas butir soal yang meliputi tiga katagori: tingkat kesulitan butir, daya pembeda, dan efektivitas distraktor.

Pelaksanaan ujicoba pertama ini sesuai dengan saran dari Brown (1995) dan Djiowandono (1995). Hakikatnya, ujicoba pertama bertujuan untuk mengetahui dan memilih butir-butir soal yang dianggap memenuhi kriteria yang diharapkan yaitu mudah, sedang, dan sulit. Proporsi butir soal dengan katagori tersebut perlu diatur secara cermat menggunakan data empiris hasil ujicoba.

Data hasil analisis koalitas butir soal dalam laporan ini menunjukkan bahwa interpretasi kualitas butir harus melibatkan tingkat kesulitan dan daya pembeda. Kedua kriteria tersebut harus mengacu pada kriteria yang bisa diterima. Untuk butir soal yang menggunakan pilihan ganda, analisis distraktor merupakan faktor penting lain yang juga harus diperhatikan. Kajian ini menunjukkan bahwa terdapat empat butir soal yang memiliki distraktor yang harus direvisi. Ini artinya, kecermatan menyusun distraktor juga merupakan hal penting dalam pengembangan butir soal. Analisis tingkat kesulitan dan daya beda menunjukkan analisis internal butir soal.

Dalam analisis validitas, diperlukan teknik yang tepat karena akan menentukan berapa kali ujicoba akan dilakukan. Kajian ini menggunakan analisis validitas kriteria karena dimaksudkan untuk mengetahui validitas eksternal. Teknik ini dianggap memiliki taraf kepercayaan yang lebih lengkap dibanding analisis internal saja karena kesesuaian dengan kemampuan peserta bisa diidentifikasi dengan cermat.

Masih terkait dengan analisis eksternal, analisis reliabilitas teknik test-retest dianggap memiliki beberapa keuntungan. Pertama, test-retest lebih praktis dilakukan karena data yang digunakan ialah data keras. Melalui dua kali tes yang diberikan pada waktu yang hampir sama terhadap peserta yang sama memberikan kemudahan bagi pengembang tes. Selain itu, analisis menggunakan rumus Product Moment lebih mudah. Namun, teknik ini bisa saja memberatkan peserta karena harus mengerjakan tes dua kali dalam waktu yang hampir sama. Kelemahan ini bisa diatasi dengan cara dipersiapkan sebelumnya dan peserta diberi waktu yang longgar.

Analisis menggunakan program SPSS ternyata lebih praktis dan menunjukkan hasil yang kurang lebih sama dengan hasil analisis manual. Ini artinya, analisis instrumen tidak perlu menggunakan teknik tertentu yang dianggap unggul dan menganggap analisis lainnya kurang baik. Faktor kepraktisan dan kemampuan peneliti merupakan pertimbangan penting untuk mengatasi persoalan ini.

Analisis SPSS dalam kajian ini menunjukkan bahwa hasil akhir yang kurang lebih sama merupakan indikasi bahwa baik manual maupun SPSS sama baiknya. Hanya saja, analisis menggunakan SPSS tidak bisa melihat lebih teliti tingkat kesulitan butir, daya pembeda, dan efektivitas distraktor. Selain itu, menggunakan SPSS kadang tidak dikuasai oleh peneliti sehingga teknik ini dianggap sulit.

Bisa disimpulkan bahwa pengembangan butir soal membaca pemahaman literal dalam kajian ini menunjukkan hasil yang akurat. Dua analisis sekaligus secara manual dan menggunakan SPSS menunjukkan bahwa kedua analisis ini tidak memberi perbedaan yang mengkhawatirkan secara metode.

SIMPULAN

Analisis pengembangan butir soal membaca pemahaman dalam kajian ini menunjukkan 24 butir valid. Analisis dilakukan menggunakan teknik manual dan teknik SPSS. Dari 40 butir yang dikembangkan berdasarkan 6 teks dan 24 butir soal yang memiliki tingkat kesulitan, daya beda, distraktor, validitas, dan reliabilitas yang bisa dipercaya.

Analisis butir di sini sengaja menggunakan dua teknik: manual dan SPSS untuk mengetahui perbedaan hasil analisis menggunakan kedua teknik tersebut. Hasil analisis menunjukkan bahwa baik manual maupun SPSS memberikan hasil yang kurang lebih

sama. Yang berbeda ialah dalam teknik manual, analisis lebih bervariasi meliputi tingkat kesulitan butir, daya beda, distraktor, validitas, dan reliabilitas. Dalam teknik SPSS, analisis mengacu pada validitas butir dan reliabilitas total saja. Masing-masing teknik memiliki keunggulan dan kelemahan. Dalam teknik manual, pelaksanaan uji coba cukup kompleks dan dilaksanakan beberapa kali. Dalam teknik SPSS program yang digunakan terbatas dan tidak semua orang menguasai.

Dalam kaitannya dengan kualitas butir yang dikembangkan di sini bisa disimpulkan bahwa kedua teknik tersebut bisa digunakan semua. Analisis butir secara meyakinkan menunjukkan bahwa ke-24 butir berkualitas bagus. Kelemahan muncul sedikit dalam bentuk distraktor pada empat butir soal yang tidak dipilih peserta sehingga perlu direvisi.

DAFTAR RUJUKAN

- Budiharso, Teguh. 2004. *Prinsip dan Strategi Pengajaran Bahasa*. Surabaya: Penerbit Lutfansyah Mediatama.
- Brown, JD. 1995. *Developing Language Test*. New Jersey: Prentice Hall Inc.
- Djaali. 2008. *Pengujian dan Pengukuran dalam Pendidikan*. Jakarta: LIPPI.
- Djiwandono, MS. 1995. *Tes dalam Pengajaran Bahasa*. Bandung: Penerbit ITB Bandung.
- Intersegmental Committee of the Academic Senates (ICAS). 2002. *Academic Literacy: A Statement of Competencies Expected of Students Entering California's Public Colleges and Universities Intersegmental Committee of the Academic Senates of the California Community Colleges, the California State University, and the University of California*. California: ICAS.
- Kim, Hyo Hyun. 2013. Needs Analysis for English for Specific Purposes Course Development for Engineering Students in Korea. *International Journal of Multimedia and Ubiquitous Engineering*, 8(6):279-288.
- Sharpe, P. 2000. *How to Prepare for the TOEFL Test*. Jakarta: Binarupa Aksara.
- Solikhah, Imroatus. 2013. English for Academic Purposes Voices: A Survey on Practices and Challenges in the State Universities of Central Java, Indonesia. *International Journal of Academic Research*, 5(4):121-125.
- Syafi'ie, Imam. 1999. *Pengajaran Membaca di Kelas-Kelas Awal Sekolah Dasar*. Pidato Pengukuhan Guru Besar. Malang: Universitas Negeri Malang.
- Yurekli, Ainur. 2012. An Analysis of Curriculum Renewal in EAP Context. *International Journal of Instruction*, 5(1):49-63.

LAMPIRAN

TES KOMPETENSI BAHASA INGGRIS

Jenis Tes	: Kompetensi Bahasa Inggris
Level	: EAP Mahasiswa Tahun Pertama
Waktu	: 90 menit

PETUNJUK

1. Tes Kompetensi Bahasa Inggris ini berjumlah 24 butir dengan cakupan: reading, dan vocabulary.
2. Jawablah setiap butir soal dalam lembar jawaban yang disediakan dengan cara memberi tanda silang (X) pada huruf A, B, C, D yang Anda anggap paling tepat.

BUTIR SOAL

Text 1

Banks are places where people can keep their money. Most people use banks to save money in their savings accounts and to pay money from their checking accounts. Today, when a person earns money from their job, their paycheck is often electronically deposited (put) into their savings or checking account. Then, he or she can pay their bills by writing checks from their checking accounts or pay online where their bills are electronically connected to their bank accounts.

Banks also give loans to people. Banks use the money that their customers deposit to lend to people to buy new houses, cars, or to start businesses among other reasons. The bank makes money from lending by charging interest. In other words, people have to pay back more than they borrowed. This amount depends on how risky the bank thinks the borrower is and how fast the loan is paid back among other things.

1. How do banks make money?
 - A. Charging interest to those they lend to
 - B. Saving their customers deposits
 - C. By having a lot of accounts
 - D. Electronically
2. What do banks NOT do?
 - A. Give loans
 - B. Charge interest
 - C. Allow people to pay bills online from their accounts
 - D. Tax people
3. What do banks NOT do?
 - A. help people get jobs
 - B. provide a place where people can pay their bills from
 - C. provide a place for people to save their money
 - D. lend money to people

4. How much "interest" do lenders have to pay?
 - A. Everyone pays the same amount of interest.
 - B. It depends on a lot of things
 - C. Most borrowers don't have to pay interest
 - D. The story doesn't say
5. How does "interest" work?
 - A. Banks require people to pay back more money than they borrowed
 - B. Banks pay people more money than they borrowed.
 - C. Banks require people to pay back the same amount they borrowed.
 - D. Banks require people to pay back money they borrowed very quickly

Text 2

The United States is a large country. It is the third largest in the whole world! It is located in a continent called North America. Parts of the United States touch three different oceans. The United States has tall mountains, wide plains, deserts, hills, rivers, lakes, volcanoes, and even rainforests! People from all over the world have come to live in the United States. There are 50 states in the United States. The newest states, Alaska and Hawaii, are not connected to the other states. Alaska is the largest state and Hawaii is a chain of islands in the Pacific Ocean.

6. Parts of the United States touch three different oceans. The words "touch" has the closest meaning to:
 - A. deal
 - B. cover
 - C. relate
 - D. linkage
7. The United States does not have...
 - A. 50 states
 - B. continents
 - C. volcanoes
 - D. mountains
8. Alaska is _____ and Hawaii is _____
 - A. a chain of islands; the largest state
 - B. the largest state; a chain of islands
 - C. not connected to the other states; the largest state.
 - D. connected to the other states; a chain of islands
9. From the text we may infer
 - A. America is chained by the Pacific Ocean
 - B. America occupies Hawaii and Alaska
 - C. America has the largest oceanic territory in the world
 - D. America has the largest states in the world

Text 3

Christopher Columbus was born in Genoa, Italy in 1451. While spending most of his early years at sea, Columbus began to believe that he could find a shortcut to the

Indies by sailing west across the Atlantic Ocean. Unfortunately, the King of Portugal refused to finance such a trip, and Columbus was forced to present his idea to the King and Queen of Spain. In 1492, King Ferdinand and Queen Isabella agreed to pay for his trip. They gave him a crew and three ships, the Nina, Pinta and Santa Maria. Columbus sailed aboard the Santa Maria.

The trip was long and hard. Many sailors grew restless and wanted to turn around. After two months at sea, land was finally sighted. The ships docked on the island of Hispaniola. Columbus named the native people he saw "Indians", because he believed he had found the shortcut he was looking for. In actuality, Columbus found North America, a brand new continent at that time. Columbus, however, couldn't be convinced. He died with the belief he had found the shortcut to the Indies. Soon, however, other explorers and nations understood the importance of his discoveries. Columbus' discoveries set the stage for the Age of Exploration, one of the most fascinating and exciting times in world history.

10. What does the text tell about?
 - A. The New Continent
 - B. Voyage to North America
 - C. Age of Exploration
 - D. Christopher Columbus

11. Which of the following statements is not true?
 - A. Columbus was born in Italy.
 - B. Columbus received three ships and a crew from the King and Queen of Spain.
 - C. Columbus found a shortcut to the Indies.
 - D. The journey across the Atlantic took two months.

12. Which of the following was NOT one of his ships?
 - A. Nina
 - B. Isabella
 - C. Pinta
 - D. Santa Maria

13. Why was Christopher Columbus very important?
 - A. He believed he found a shortcut to the Indies.
 - B. He first used the word "Indians".
 - C. He discovered a whole new continent.
 - D. He was one of the bravest explorers of all time.

Text 4

In ancient times, the king placed a boulder on a roadway. Then he hid and watched to see if anyone would remove the huge rock. Some of the king's wealthiest merchants and courtiers came by and simply walked around it.

Many loudly blamed the king for not keeping the roads clear, but none did anything to get the big stone out of the way. Then a peasant came along carrying a load of vegetables. On approaching the boulder, the peasant laid down his burden and tried to move the stone to the side of the road. After lots of pushing and straining, he finally

succeeded. As the peasant picked his load of vegetables, he noticed a purse lying on the road where the boulder had been.

The purse contained a lot of gold coins and a note from the king indicating that the gold was for the person who removed the boulder from the roadway. The peasant learned what many others never understand.

14. Why did many people blame the king?
 - A. He did not make good roads for them
 - B. He loved to hide behind the rock
 - C. He showed no care on their roads.
 - D. He did not keep the road clear.

15. What does the last paragraph tell us?
 - A. The roads that people built with the king.
 - B. The person who liked keeping the roads well.
 - C. The purpose of putting the boulder and purse.
 - D. The way the peasant removed the boulder on the road.

16. From the story, we know that...
 - A. Many people liked the peasant
 - B. The peasant was a good man
 - C. The boulder was really small
 - D. The king was very stingy

17. What can we learn from the story?
 - A. A problem may be a blessing in disguise
 - B. There so many problems in life
 - C. An obstacle may make us stronger
 - D. We have to avoid any problem in our life.

Text 5

In 1854, Louis Pasteur was appointed a professor of chemistry at the University of Lille. He demonstrated that organism such as bacteria were responsible for souring wine, beer even milk. He then invented a process where bacteria could be removed by boiling and then cooling the liquid. Today the process is known as pasteurization. He completed the first test on April 20, 1862.

In 1865, Pasteur helped to save the silk industry. He proved that microbes were attacking healthy silkworm eggs, causing the diseases, and the disease can be cured by eliminating the microbes. Pasteur first vaccine discovery was in 1879, with a disease called chicken pox. After accidentally exposing chicken to the attenuated form, he demonstrated that they became resistant to the actual virus. Pasteur went on to extend

his germ theory to develop prototype of vaccine for disease such as, anthrax, cholera, TB and smallpox.

18. What happened to Pasteur in 1879?

- A. He found his first vaccine.
- B. He was appointed as chemistry professor.
- C. He developed the vaccine prototype.
- D. He invited pasteurization process.

19. From the text, we can conclude that...

- A. Louis Pasteur is a method of processing milk.
- B. Louis Pasteur found no cure for anthrax and cholera.
- C. Louis Pasteur had his own silk industry..
- D. Louis Pasteur invented a method called pasteurization

20. What is the main idea of the last paragraph?

- A. Louis Pasteur completed the test well
- B. Louis Pasteur demonstrated his theory
- C. Louis Pasteur became a professor
- D. Louis Pasteur discovered vaccination

21. "He then invented..." (first paragraph).

What does the word "invented" mean?

- A. looked for
- B. used
- C. created
- D. thought about

Text 6

Vacation Time Inc.

Specializing in discount vacations.

We make all your reservations for you.

We guarantee:

*the lowest fare on all major airline routes.

**the most economical hotel rates.

***the least expensive rental cars.

Why pay more?

Call us today.

22. What kind of a business is Vacation Time Inc?

- A. a travel agency
- B. an airline
- C. a hotel
- D. a vacation resort

23. Which is the one service NOT mentioned?
- | | |
|-------------------------|---------------------------|
| A. airline reservations | C. rental car reservation |
| B. hotel reservations | D. cruise reservation |
24. What is the core of Vacation Time's business?
- | | |
|--------------------------|------------------------|
| A. high-end luxury goods | C. limousine pickup |
| B. travel bargains | D. self-arranged tours |

