



PROCEEDING



THE INTERNATIONAL CONFERENCE ON LANGUAGE TEACHING AND ASSESSMENT

"Language Teaching and Assessment
for the 21st Century: Current Issues and Trends"

Syahida Inn, August 22-23, 2017



Pusat Pengembangan Bahasa
Universitas Islam Negeri Syarif Hidayatullah
Jakarta



Pusat Pengembangan Bahasa
Universitas Islam Negeri Syarif Hidayatullah
Jakarta



9 772581 137000

PROCEEDING

INTERNATIONAL CONFERENCE ON LANGUAGE TEACHING AND ASSESSMENT

"Language Teaching and Assessment
for the 21st Century : Current Issues and Trends"

Editors:

Siti Nurul Azkiyah
Teguh Khaerudin
Yenny Rahmawati
Mukhsin Nawawi
Azkia Muharom Albantani

Design and Layout:

Fakhrul Arifin
Yazid Hady

ISSN : 2581-1371

Publisher:

Pusat Pengembangan Bahasa
UIN Syarif Hidayatullah Jakarta
Jl Ir H Juanda No 95 Tangerang Selatan 15419
Phone. 021 744 3328

Size : 17,6 cm x 25 cm
x + 179 pages

Copyright 2017 © Pusat Pengembangan Bahasa UIN Jakarta

CHAPTER 2 : ARABIC

THE EFFECTIVENESS OF IMPLEMENTATION TAMYIZ METHOD TO IMPROVE 65
SANTRIS' ABILITY IN READING KITAB KUNING AT MODERN ISLAMIC BOARDING
SCHOOL AL-JUNAIDIYAH BIRU BONE REGENCY, SOUTH SULAWESI

Muhammad Sultan
STAIN Watampone

مدرج تعليم اللغة العربية المبني على وظيفتها الدينية والتواصلية (مدرج مقترح للمدارس 76
الإعدادية والثانوية بجمهورية إندونيسيا)
أحمد فؤاد أنسي
جامعة مالانج الحكومية

أسلوب التعلم التعاوني بظراظ تنسيمي الطلبة إلى المجموعة في تعليم قراءة النصوص العربية لغة 83
ثانية

Asrina
UIN Imam Bonjol Padang

نموذج التعلم الإلكتروني عبر الواتساب في تدريس اللغة العربية 98

Raswan
UIN Syaif Hidayatullah Jakarta

طريقة تعليم مهارات اللغوية لدى الطلاب في الجامعة 114

Siti Solichah
UIN Sultan Maulana Hasanuddin

تعليم اللغة العربية باستخدام الحاسوب (Macromedia Flash) في مادة الإملاء لتنمية مهارة 121
الكتابة (دراسة تجريبية في جامعة الإسلامية الحكومية سمارندا)
لولوك حنيبا عسافا، عبد الرزاق، مراحو
جامعة الإسلامية الحكومية سمارندا

✓ ANALISIS BUTIR TES TOAFL (TEST OF ARABIC AS A FOREIGN LANGUAGE) UIN 131
MALIKI MALANG
✓ Syaiful Mustofa
UIN Maulana Malik Ibrahim Malang

CHAPTER 3 : INDONESIAN

PERANCANGAN BUKU KOMIK BAHASA INDONESIA BAGI PENUTUR ASING (BIPA) 141
Randi Ramliyana, Avini Nurazhimah Arfa
Universitas Indraprasta PGRI

ANALISIS BUTIR TES SOAL *TOAFL*

UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG

Syaiful Mustofa

Email: saifulmustofa@pba.uin-malang.ac.id

A. Latar Belakang

Berdasarkan rekomendasi hasil rapat kerja Universitas Islam Negeri Maulana Malik Ibrahim Malang tentang sosialisasi program dan anggaran tahun 2016 dan perancangan kegiatan tahun 2017 di Prigen Pasuruan pada tanggal 19-21 Februari 2016 lalu, menyebutkan bahwa tes *TOAFL* akan diberlakukan untuk calon wisudawan mahasiswa S1, S2 dan S3 di lingkungan Universitas Islam Negeri Maulana Malik Ibrahim Malang. Diantara tujuan diselenggarakannya tes *TOAFL* ini adalah untuk mengetahui sejauh mana tingkat kebahasaan yang dimiliki oleh mahasiswa setelah memperoleh pendidikan dan pengajaran dalam waktu yang telah ditentukan. Disamping itu, penyelenggaraan tes *TOAFL* ini akan menjadi salah satu SKPI (Sertifikat Kompetensi Pendamping Ijazah) yang digalakkan Pemerintah akhir-akhir ini.

Perguruan Tinggi Keagamaan Islam, baik itu UIN, IAIN, STAIN menyepakati bahwa *TOAFL* adalah salah satu alat tes yang bisa digunakan untuk mengukur kemampuan berbahasa para mahasiswa. Standarisasi penilaian dalam *TOAFL* pun telah disepakati oleh Perguruan Tinggi Keagamaan Islam itu sendiri.

Soal yang diujikan dalam *TOAFL* tidak hanya didasarkan pada kebutuhan akademik saja, namun juga didasarkan pada isu-isu terkini dalam bahasa Arab, baik melalui internet, Koran arab, majalah arab dan lain sebagainya. Namun sangat disayangkan sekali bahwa tes semacam *TOAFL* ini belum dikembangkan oleh bangsa Arab sendiri. Dalam penelusuran penulis, pencetus tes ini orang Indonesia, yaitu Muhibb Abdul Wahhab dari UIN Syarif Hidayatullah. Kendati demikian, kepemilikan sertifikat *TOAFL* ini sudah menjadi syarat mutlak dalam melanjutkan studi pascasarjana di Perguruan Tinggi Keagamaan Islam di seluruh Indonesia. Tidak hanya sampai disitu, sertifikat *TOAFL* ini juga digunakan untuk seleksi beasiswa doktoral dan tes menjadi petugas haji dll.

Tes *TOAFL* dilaksanakan di beberapa kampus dengan butir soal tesnya telah terdaftar di HAKI (hak kekayaan intelektual) dan mempunyai lisensi tertentu termasuk

yang akan dilakukan oleh Pusat Pengembangan Bahasa (PPB) Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang . Soal-soal yang ada di dalam TOAFL yang mau didaftarkan ke HAKI, kiranya peneliti perlu melakukan analisis dan evaluasi yang mendalam sehingga dapat dihasilkan sebuah soal TOAFL yang benar-benar memenuhi syarat pembuatan soal yang baik.

Disamping tes TOAFL sebagai sebuah kebutuhan primer dalam mengembangkan dunia akademis, ia juga merupakan kebutuhan sekunder yang perlu dimiliki oleh seluruh Perguruan Tinggi Keagamaan Islam baik negeri maupun swasta. Tes ini akan dapat dijadikan ukuran penting untuk mengukur kemampuan berbahasa Arab calon lulusannya.

Pusat Pengembangan Bahasa (PPB) Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang telah mengembangkan tes TOAFL sejak 2005, dan telah *upgrade* soal-soalnya pada tahun 2016. Sama pada tahun 2005, tes TOAFL yang dikembangkan pada 2016 pun belum didaftarkan ke HAKI, akan tetapi akan berencana mendaftarkan ke HAKI setelah peneliti melakukan evaluasi dan analisis tes butir soal TOAFL yang telah dikembangkan. Beberapa butir soal yang dirasa tidak valid, akan diperbaiki sampai menjadi soal yang benar-benar valid. Atas dasar itulah, bahwa penelitian ini layak dan *urgen* untuk dilakukan.

Obyek penelitian ini adalah tes TOAFL yang dikembangkan Pusat Pengembangan Bahasa (PPB) UIN Maliki Malang tahun 2016. Hal ini didasarkan atas pertimbangan bahwa tes TOAFL UIN Maliki Malang menjadi rujukan mahasiswa seluruh Indonesia untuk mendapatkan sertifikat. Butir tes TOAFL UIN Maliki Malang sudah di *upgrade* pada bulan januari dan pebruari 2016 lalu, akan tetapi butir tes tersebut belum diukur validitas, realibilitas, tingkat daya beda, dan tingkat kesukaran sehingga belum diketahui standart validitas, realibilitas, tingkat daya beda dan tingkat kesukaran butir tes TOAFL tersebut.

B. Kajian Pustaka

1. Konsep Analisis Butir Tes

Dalam ilmu evaluasi pendidikan diketahui bahwa analisis butir soal merupakan instrumen yang perlu dilakukan agar seorang guru memperoleh soal yang benar benar baik dan benar. Tujuan utama dalam analisis butir soal ini adalah (a) guru/dosen akan

berupa dengan rumus yang ada menentukan validitas dan reliabilitas soal dan (b) guru/dosen mampu menganalisis butir soal.

Adapun analisis butir soal tes yang telah dijawab oleh mahasiswa/siswa dalam suatu ujian berfungsi untuk (a) mengetahui mana saja soal yang dianggap sulit oleh mahasiswa/siswa dan mengetahui mahasiswa/siswa mana saja yang mengalami kendala dalam mengerjakan soal tersebut, sehingga guru/dosen akan melakukan pembimbingan khusus kepada mahasiswa/siswa yang mengalami kesulitan tersebut, dan (b) untuk mendeteksi lebih mendalam soal-soal yang memiliki daya beda, yang kemudian akan digunakan landasan revisi soal untuk menyiapkan soal yang lebih baik dan benar. Dari pernyataan diatas, peneliti dapat mengambil kesimpulan bahwa sebuah analisis soal memiliki empat hal utama yaitu:

- (a) Bagaimanakah taraf penguasaan mahasiswa/siswa terhadap setiap butir tes soal
- (b) Apakah soal tersebut memiliki validitas dan realibilitas tinggi atautakah tidak
- (c) Apakah soal tersebut memiliki daya pembeda yang baik atautakah tidak
- (d) Sejauh mana tes tersebut dalam mengukur kemampuan berbahasa mahasiswa/siswa

2. Tes sebagai Instrumen Evaluasi

Tes merupakan salah satu instrument evaluasi yang berfungsi untuk mendeteksi kemampuan peserta didik dalam berbahasa sesuai dengan level belajarnya. Untuk menjadi peserta tes, seorang peserta wajib memiliki syarat-syarat tertentu yang harus dipenuhinya.

Dalam bahasa Arab, tes dikenal dengan nama *ikhtibaar*. Dan dalam ilmu evaluasi pendidikan bahasa Arab, TOAFL merupakan salah satu tes yang cocok digunakan untuk mengukur level kemampuan berbahasa peserta didik. Tes TOAFL ini dirancang sedemikian rupa untuk tujuan dan sasaran belajar peserta didik itu sendiri. Standar penilaian pada TOAFL ini pun diberitahukan secara jelas kepada khalayak umum. Baik dosen maupun mahasiswa dapat mengetahui macam-macam level dalam standar penskoran TOAFL itu sendiri.

Dosen selaku pendidik harus mengetahui sejauh mana tingkat pemahaman dan level kebahasaan yang dimiliki oleh peserta didik. Dengan demikian, dosen akan memperoleh gambaran secara sempurna tentang level kebahasaan mahasiswanya. Ini

diupayakan dalam rangka tugas dosen untuk melakukan evaluasi dan penilaian terhadap pemerolehan hasil belajar mahasiswanya.

Tes TOAFL merupakan sejumlah prosedur dan instrument yang didesain secara sistematis, digunakan oleh seseorang atau lembaga untuk menguji kemampuan berbahasa seseorang dengan maksud dan tujuan tertentu. Sehingga petunjuk dalam pengerjaannya pun disusun secara jelas dan tidak memungkinkan untuk salah tafsir. Beberapa contoh petunjuk dalam mengerjakan tes dalam TOAFL adalah melingkari pilihan jawaban yang benar, memilih diantara dua jawaban yang benar/salah, mengisi kata-kata yang kosong pada soal, menjelaskan gambar atau tabel dengan pemahaman peserta tes didik.

3. Bentuk Tes Bahasa Arab

Bentuk tes bahasa sangatlah banyak bergantung pada kepentingan, tujuan, ruang lingkup dan cara pengerjaannya. Dari segi tujuannya tes bahasa dibagi menjadi tiga yaitu Tes prestasi, tes profisiensi dan tes kesiapan berbahasa/ tes prediksi.

Gambaran umum tes pemerolehan bahasa adalah sebuah tes yang digunakan untuk menguji sejauh mana peserta didik dalam memperoleh pengalaman pendidikan dan pengajaran dalam waktu tertentu. Tes ini juga didasarkan pada kurikulum atau bahan ajar yang digunakan, apakah ada hubungan timbal balik yang baik, ataukah terjadi perbedaan yang mandalam diantara keduanya. Jika kurikulum dan bahan aja sudah benar, maka instrument tesnya lah yang perlu di revisi dan sebaliknya. Termasuk dalam tes pemerolehan bahasa ini adalah Tes Tengah Semester dan Tes Akhir Semester.

Adapun Tes profisiensi bertujuan untuk mengetahui level berbahasa peserta didik. Ujian ini tidak terkait dengan buku ajar yang ia gunakan, kurikulum yang digunakan kampus atau berapa lama peserta tersebut dalam belajar bahasa. Tes ini lebih mengarah kepada ujian kemampuan berbahasa secara umum. Termasuk dalam tes ini adalah tes TOAFL dan TOEFL.

Macam lain dari tes adalah tes kesiapan berbahasa. Tes ini bertujuan untuk mengetahui tingkat kesiapan peserta untuk belajar bahasa kedua dan digunakan juga untuk memprediksi kemajuan peserta didik. Terkadang tes ini juga digunakan untuk penempatan kelas pada suatu program pendidikan. Tes ini terdiri dari tes *audio visual* dan *taraakib wa al-mahaarat al-luhowiyah* untuk benar-benar memperoleh tempat yang

sesuai dengan level yang dimiliki peserta didik.

Sedangkan bila dilihat dari segi pembuatnya, tes terbagi menjadi dua yaitu tes tenaga pengajar dan tes standar. Tes tenaga pengajar merupakan tes yang dibuat oleh seorang guru/dosen bertujuan untuk menguji kemampuan bahasa mahasiswa/siswanya. Tes ini akan digunakan oleh seorang guru/dosen untuk menentukan strategi apa yang akan digunakan untuk pembelajaran setelahnya. Tes standar merupakan tes yang dibuat oleh sebuah lembaga/institusi yang digunakan untuk skala luas (sekolah se kota Malang, atau se Jawa Timur), tidak hanya untuk skala sempit (sekolah/universitas) tertentu.

Sebuah tes bila dilihat dari metode penskorannya terbagi menjadi dua, yaitu tes subjektif (esai) dan tes obyektif (pilihan ganda). Tes esai ini disebut tes subyektif karena jawaban peserta didik yang diberikan pendidik bersifat bebas, namun nilai yang diberikan oleh seorang pendidik bersifat subjektif. Sedangkan tes obyektif merupakan soal yang dibuat oleh pendidik dan pendidik memberikan beberapa pilihan ganda dengan keseragaman data. Tes ini disebut tes obyektif dikarenakan tes ini bersifat tertutup dan tidak membuka pilihan jawaban yang sudah ditentukan, misal pilihan A,B,C,D,E maka salah satu diantara 5 jawaban itu hanya akan ada satu jawaban yang benar.

Dalam tes obyektif terdapat beberapa type soal yakni tes pilihan ganda, tes melengkapi isian, dengan jawaban yang bersifat tertutup, tes pilihan benar salah tes mencari pasangan.

Sedangkan bila dilihat dari cara dan bentuk pengujiannya, tes dibagi menjadi dua yaitu tes tulis (*al-iktibar al-tahriri*) dan tes lisan (*al-ikhtibar al-syafawi*). Tes lisan adalah model tes yang disajikan dalam tulisan/kata kata sedangkan tes lisan adalah tes yang diberikan secara lisan, seperti ujian berbicara bahasa arab, ujian pidato bahasa arab, ujian meneruskan puisi berbahasa Arab dll. Sedangkan tes tulis adalah soal yang diujikan melalui tulisan. Model tes tulis sangat cocok untuk menguji ilmu cabang bahasa Arab seperti tes nahwu, tes terjemah, tes mengarang teks Arab dan lain sebagainya.

Semua model tes diatas haruslah memenuhi unsur nsur sebagai beikut yaitu:

- (a) Butir soal harus sesuai dengan tujuan khusus yang ditetapkan. Misal jika butir soal yang diujikan adalah *mufrodāt* maka hendaknya soal tersebut tidak bias dengan unsur sintaksis (*shorof*) dan morfologi (*nahwu*) bahasa Arab.
- (b) Butir soal yang dibuat harus memperhatikan tingkat kemampuan peserta didik, terutama bila yang diujikan adalah tes pemerolehan bahasa.
- (c) Penyusunan tes hendaknya disertai dengan petunjuk jelas, baik sisi waktu yang digunakan, tempat menjawabnya dan bagaimana cara menentukan jawabannya.
- (d) Redaksi soal tidak boleh bersayap atau multiinterpretasi, tidak diskriminatif dan terukur dengan baik
- (e) Waktu yang ditentukan harus sebanding dengan jumlah soal dan tingkat kesulitan yang dihadapi peserta ujian
- (f) Skoring peniaian bersifat objektif berdasarkan rumus yang telah ditetapkan di awal, dan sangat jauh dari subyektivitas penilai.

4. TOAFL

Test of Arabic as Foreign Language atau biasa disebut dengan nama TOAFL merupakan nama yang digunakan untuk sebuah tes bahasa Asing untuk bukan penutur bahasa Asing itu sendiri. Nama TOAFL terilhami dari nama tes TOEFL (*Test of English as a Foreign Language*). Penyebutan nama TOAFL ini dimaksudkan untuk mudah diucapkan dan mudah dipahami, bahwa ia adalah nama untuk sebuah jenis tes resmi kepemilikan level berbahasa Asing.

TOAFL berdiriatas sebuah alas an yang jelas dan mendasar yakni keinginan PTKIN untuk mengukur secara jelas kemampuan berbahasa para lulusannya. Penyusunan soal TOAFL tidak hanya murni tentang bahasa Arab saja, tapi juga unsur ilmu-ilmu keislaman seperti fiqih, tafsir, hadits dll. Hal ini sejalan dengan misi dari tes TOAFL itu sendiri yaitu bahwa tes TOAFL membantu mewujudkan bahasa Arab sebagai bahasa resmi dalam mempelajari Islam dan ilmu Sains Teknologi.

Sejak beberapa tahun terakhir ini (sejak 2005an) tes TOAFL telah menjadi salah satu syarat masuk di perguruan tinggi Islam dalam dan luar negeri. Bahkan untuk seleksi masuk program beasiswa dalam dan luar negeri telah ditetapkan kepemilikan

kemampuan berbahasa TOAFL dengan skor yang berbeda. Misal untuk lulus S1 harus memiliki skor minimal 450, untuk S2 minimal 475 dan untuk S3 minimal 500, untuk seleksi beasiswa doctoral dalam negeri 500 dan luar negeri 525. Skor yang ada ini adalah skor minimal, dan apabila di atas nilai skor tersebut maka tidak akan menjadi masalah, bahkan dianjurkan.

5. Mengukur Butir Tes TOAFL

a) Validitas

Validitas adalah sebuah soal yang apabila terukur dan sesuai antara apa yang diukur dengan tes yang digunakan. Validitas dalam soal TOAFL dimaksudkan ialah sebuah tes yang kesesuaian antara soal tes dengan apa yang diukur dalam tes TOAFL itu sendiri.

Ada yang membedakan validitas menjadi tiga bagian yaitu validitas kriteria (*criteria validity*), validitas isi (*content validity*) dan validitas konstruk (*construct validity*). Ada juga ilmuwan lain yang membagi validitas menjadi empat yaitu validitas konkuren (*concurent validity*), validitas prediktif (*predictive validity*), validitas konstruk (*construct validity*) dan validitas isi (*content validity*).

Validitas konstruk merupakan hasil dari sebuah tes dapat diperoleh pembuktian validitas konstruk. Dalam tes kemampuan menulis (*kitaabah*) misalnya, benar-benar mencerminkan antara jenis dan rincian kemampuan menulis bahasa Arab yang sama dengan jenis dan rincian kompetensi dan indikator yang diperlukan untuk memahami kemampuan menulis bahasa Arab itu sendiri.

Validitas ini wajib menunjukkan kesesuaian isi antara tes yang akan digunakan sebagai instrument alat tes dan kemampuan yang diukur dengan nya. Seperti sebuah tes yang digunakan untuk menguji kemampuan *istimak* haruslah benar-benar menguji tentang kemampuan peserta didik dalam mendengarkan teks bahasa Arab, bukan yang lain. Artinya, sebuah tes yang digunakan untuk maharah tertentu haruslah sesuai dengan maharah yang diuji tersebut.

Validitas kriteria didasarkan atas kesesuaian antara hasil suatu tes dengan hasil tes lain. Apabila ternyata soal yang digunakan memiliki kriteria yang baik, maka hasil yang didapatkan kurang lebih sama. Dalam buku lain, disebutkan bahwa validitas kriteria ini disebut juga sebagai validitas waktu.

Bagi seorang ilmuwan di bidang akademik akan paham bahwa validitas sangat berkaitan dengan kelayakan penafsiran sebuah skor dengan kemampuan yang dimiliki oleh orang yang memiliki skor tersebut. Seorang ilmuwan seperti Gronlund, Popham dan Nurgiyanto menegaskan validitas dalam sebuah tes digambarkan sebagai berikut:

Tabel 1. Tiga Pendekatan Validasi Tes

Jenis Pendekatan	Prosedur	Keterangan
Bukti berdasarkan isi (<i>Content-Related Evidence</i>)	Perbandingan butir-butir tes dengan deskripsi spesifikasi tes	Sejauh mana sampel tes mewakili kompetensi ranah yang diukur
Bukti berdasarkan kriteria (<i>Criterion-Related Evidence</i>)	Perbandingan skor hasil tes dengan skor hasil yang berada di dalam skor tersebut	Kesesuaian antara tes dan hasil yang ada pada skor tes
Bukti berdasarkan konstruk (<i>Construct-Related Evidence</i>)	Penetapan makna dibalik skor yang dihasilkan peserta tes, dengan memperhatikan berbagai macam aspek dan faktor yang mempengaruhi tampilan tes itu	Sejauh mana tes tersebut memiliki standart yang sama bila diujikan dengan orang lain dan kemampuan yang setara

b) Reliabilitas

Reliabilitas merupakan salah satu ciri tes dalam pembuatan tes yang baik. Reliabilitas ini dimaksudkan sebagai standar ukur yang tetap, ajek, dan tidak berubah-ubah. Dengan kata lain, reliabilitas tidak hanya digunakan sebagai alat ukur saja akan tetapi jauh dari tujuan itu yaitu ia merupakan instrument yang stabil, stagnan, ajek dan tidak mudah berubah-ubah. Dengan ciri keajegan itu, sebuah tes akan menghasilkan skor yang sama apabila diujikan dengan orang yang berbeda.

Konsistensi sebuah pengukuran dapat diketahui dengan reliabilitas. Soal dikatakan reliable apabila soal tersebut memiliki derajat kepercayaan dan andal. Hasil yang sama pada sebuah tes yang telah diujikan kepada subjek lebih dari satu kali dan dalam waktu

yang berbeda disebut juga sebagai soal yang memiliki konsistensi yang tinggi. Instrument dianggap ajeg apabila memiliki reliabel tinggi.

Sebuah instrument tes dikatakan reliabel atau handal apabila memiliki rasio yang terukur antara varian skor yang diperoleh (*observed score*) dengan hasil skor murni (*true skor*). Dengan kata lain bahwa semakin besar nilai variable pada skor yang diperoleh peserta tes menunjukkan bahwa tes tersebut dinilai baik dan sebaliknya. Apabila sebuah tes TOAFL memiliki variabel tinggi, maka tes tersebut layak disebut sebagai instrument tes yang baik.

c) Daya Beda

Daya beda merupakan salah satu ciri sebuah tes yang digunakan untuk membedakan antara peserta dengan kemampuan tinggi (pintar) dengan peserta dengan kemampuan rendah (kurang pandai). Sedangkan derajat kesulitan adalah sebuah nilai yang digunakan untuk mengklasifikasikan mana saja siswa yang bisa melewati soal-soal sukar dan mana saja siswa yang belum mampu dalam melewati soal sukar tersebut.

Daya pembeda suatu butir soal dapat dimaknai sebagai alat yang digunakan mengukur sejauh mana soal itu bisa membedakan kelompok testi yang pandai dan kelompok testi yang lemah. Rumus yang dipakai adalah sebagai berikut:

$$D = P_H - P_L$$

D : Indeks Daya Pembeda

P_H : peserta tes kelompok atas yang menjawab benar butir tes

P_L : peserta tes kelompok bawah yang menjawab benar butir tes

Adapun tabel indeks daya pembeda yang digunakan sebagai berikut:

Tabel 2. Penafsiran Indeks Daya Pembeda

Indeks Daya Pembeda	Kategori Butir Daya Pembeda
$0,40 < D$	sangat baik
$0,30 < D \leq 0,40$	baik
$0,20 < D \leq 0,30$	cukup
$D \leq 0,20$	jelek

d) Tingkat Kesukaran

Sebuah soal dikatakan ideal apabila soal tersebut benar-benar sesuai dengan kemampuan peserta tes. Artinya, apabila ada soal yang hanya bisa dikerjakan oleh segelintir orang maknanya soal tersebut bukanlah soal yang baik, karena merupakan soal yang sukar dan hanya mampu dikerjakan oleh kelompok atas (*upper group*). Pun sama halnya apabila sebuah soal bisa dikerjakan semuanya oleh peserta tes, maka sesungguhnya soal itu juga tidak baik, karena terlalu mudah, dan soal jenis ini tidak mampu membedakan antara peserta tes yang pandai dan yang tidak pandai. Angka-angka kuantitatif dalam skala tingkat kesulitan (TK) dirumuskan oleh Thorndike dan Hagen dalam bukunya (*Measurement and Evaluation in Psychology and Education*) sebagai berikut:

Tabel 3. Tingkat Kesukaran

TINGKAT KESUKARAN (TK) atau <i>p</i>		
0,71 - 1,00	:	mudah (Md)
0,30 - 0,70	:	sedang (Sd)
0,00 - 0,29	:	sukar (Sk)

Sedangkan rumus yang digunakan mencari derajat kesukaran sebuah soal adalah:

TK = PB : JP		
TK	:	Tingkat kesukaran soal
PB	:	Peserta yang menjawab benar
JP	:	Jumlah seluruh peserta

Setelah soal TOAFL tersebut dilakukan analisis dengan rumus diatas, peneliti akan mengetahui ternyata dalam soal yang itu terbagi menjadi 3 macam yaitu

- Soal sukar, yaitu soal yang memiliki tingkat kesukaran berkisar 0,00-0,29
- Soal sedang yaitu soal yang memiliki tingkat kesukaran berkisar 0,30-0,70
- Soal mudah, yaitu soal yang memiliki tingkat kesukaran berkisar 0,70-1,00

Tabel 4. Indeks Soal

Rentang indeks soal	<i>FEED BACK</i> YANG DILAKUKAN
Sukar (0,00 – 0,29)	<ul style="list-style-type: none"> ▪ Soal disimpan sementara atau sekalian didelete karena terlalu sukar ▪ Soal di kaji lagi(redaksi soal,bentuk pertanyaan, dll) ▪ Soal digunakan untuk tes level tinggi (<i>mutaqoddim</i>)
Sedang (0,30 – 0,70)	<ul style="list-style-type: none"> ▪ Soal dipakai dan dipertahankan tingkat daya bedanya ▪ Soal disimpan untuk digunakan kembali
Mudah (0,71 – 1,00)	<ul style="list-style-type: none"> ▪ Soal disimpan sementara atau sekalian didelete karena terlalu mudah ▪ Soal dikaji ulang penyebabnya mudah dijawab ▪ Soal digunakan untuk tes biasa

C. Hasil Penelitian

Peneliti menganalisis butir soal Maharah Istima' Tes TOAFL UIN Maliki Malang dan memperoleh hasil sebagai berikut:

1. Reliabilitas

Temuan pertama dapat diinformasikan bahwa unsur temuan tentang reliabilitas butir tes TOAFL UIN Maulana Malik Ibrahim Malang. Jika kembali pada teori reliabilitas pada soal dapat disimpulkan bahwa soal berkategori baik apabila lebih dari atau sama dengan 0,70 dan berkategori jelek bila kurang dari 0,70.

2. Daya Beda

Rumus kriteria daya beda adalah negatif (-), dan bila kurang dari 0,20 maka dikategorikan soal berdaya beda lemah, apabila antara 0,20-0,39 disebut soal berdaya beda cukup, apabila antara 0,40-0,69 soal berdaya beda baik dan apabila 0,70-1,00 maka berdaya beda baik sekali.

Tabel 5. Daya Beda

NO	DAYA PEMBEDA	MAKNA
1.	$0 < 0,20$	Lemah
2.	$0,20 - 0,39$	Cukup
3.	$0,40 - 0,69$	Baik
4.	$0,70 - 1,00$	Baik Sekali

Tabel 6. Prosentase Daya Beda

NO	DAYA PEMBEDA	KETERANGAN	JUMLAH	PERSENTASE
1.	$0 < 0,20$	Lemah	13	26%
2.	$0,20 - 0,39$	Cukup	9	18%
3.	$0,40 - 0,69$	Baik	15	30%
4.	$0,70 - 1,00$	Baik Sekali	13	26%

3. Tingkat Kesulitan

Tingkat kesulitan soal adalah mengkaji soal-soal tes dari segi kesulitannya sehingga dapat diperoleh soal-soal mana yang termasuk mudah, sedang, dan sulit. Kriteria tingkat kesulitan $0,00 - 0,30$ termasuk soal kategori sulit, $0,31 - 0,70$ termasuk soal kategori sedang, $0,71 - 1,00$ termasuk soal kategori mudah. Berdasarkan hasil penelitian menunjukkan bahwa tidak ada soal yang termasuk kategori sulit ada 5 soal (10%), soal yang termasuk kategori sedang ada 26 soal (52%), dan soal yang termasuk ke dalam kategori mudah ada 19 soal (38%).

Tabel 7. Tingkat Kesulitan

NO	TINGKAT KESULITAN	JUMLAH	PERSENTASE
1.	$0,71 - 1,00$ (Mudah)	19	38%
2.	$0,31 - 0,70$ (Sedang)	26	52%
3.	$0,00 - 0,30$ (Sulit)	5	10%

Adapun soal maharah istima' TOAFL UIN Maulana Malik Ibrahim Malang secara umum mempunyai daya beda yang cukup baik, yaitu dengan 13 (26%) soal yang mempunyai daya beda yang rendah atau lemah, 9 (18%) soal yang tergolong mempunyai daya beda yang cukup, 15 (30%) soal mempunyai daya beda yang baik, dan 13 (26%) soal yang mempunyai daya beda yang baik sekali. Kriteria untuk daya pembeda adalah apabila negatif (-) berarti tidak ada daya pembeda, < 0,20 dikategorikan daya beda lemah, 0,20 – 0,39 kategori daya beda cukup, 0,40 – 0,69 kategori daya beda baik, 0,70 – 1,00 kategori daya beda baik sekali. Sehingga dapat kita simpulkan bahwa soal maharah istima' TOAFL UIN Maulana Malik Ibrahim Malang mempunyai daya beda yang baik.

Tingkat kesulitan soal diperoleh melalui analisis butir soal yang kemudian dimasukkan ke dalam rumus dan mengkajinya secara mendalam. Dari sana ditemukan data soal-soal yang berkategori kesulitan mudah, sedang, dan sulit. . Berdasarkan hasil penelitian, peneliti mengetahui bahwa ada 5 soal berkategori sulit (5%), dan 26 soal berkategori sedang (52%), dan 19 soal berkategori mudah (38%).

Dengan demikian bahwa soal maharah istima' TOAFL UIN Maulana Malik Ibrahim Malang ada 33 (66%) soal yang direkomendasikan untuk digunakan dan disimpan di bank soal, dan 2 (4%) soal yang disarankan untuk direvisi atau diperbaiki sehingga bisa menjadi soal yang baik dan mempunyai reliabilitas dan daya beda yang baik, kemudian ada 15 (30%) soal yang masih belum mempunyai ciri-ciri atau kriteria soal yang baik, sehingga peneliti menyarankan untuk diganti soal baru yang lebih baik dan mempunyai reliabilitas dan daya beda yang baik.

Tabel 8. Persentase Status Soal

NO	STATUS SOAL	JUMLAH	PERSENTASE
1.	Diterima	33	66%
2.	Direvisi	2	4%
3.	Ditolak	15	30%

D. Penutup

Setelah peneliti melakukan analisis butir soal, kesulitan soal, reliabilitas soal, daya pembeda soal, peneliti menyimpulkan hal-hal sebagai berikut:

1. Soal-soal yang ada pada tes TOAFL UIN Maulana Malik Ibrahim Malang pada aspek tes *mahaarah istimaa'* dapat dikategorikan sebagai soal "baik" diindikatori oleh pedoman evaluasi penskoran tes yang soal baik
2. Hasil penelitian terhadap analisis reliabilitas soal tes TOAFL UIN Maulana Malik Ibrahim Malang khususnya *mahaarah istimaa'* memiliki reliabilitas tingkat tinggi, dijelawantahkan dengan nilai reliabel $r_{11} \geq 0,70$ yaitu 0,799
3. Melalui SPSS dapat diperoleh data sebagai berikut soal *mahaarah istimaa'* TOAFL UIN Maulana Malik Ibrahim Malang sebanyak 13 soal memiliki daya beda lemah, 9 soal memiliki daya beda cukup, 13 soal berdaya beda baik. Dan secara umum, soal *mahaarah istimaa'* TOAFL UIN Maulana Malik Ibrahim Malang tergolong soal berdaya beda baik.
4. Peneliti melakukan analisis soal dan memperoleh data bahwa ada 5 soal (10%) merupakan soal berkategori sulit, 26 soal (52%) berkatageoei sedang dan 19 soal (38%) berkategori mudah

Setelah menyelesaikan penelitian ini serta menyimpulkannya, peneliti menyarankan beberapa hal untuk dilaksanakan dan dilanjutkan, yaitu :

- a. Peneliti melakukan pencatatan langsung soal-soal yang memiliki daya pembeda berkategori "baik" untuk selanjutnya digunakan pada tes mendatang, walaupun menggani yang diganti hanyalah kosakata yang sepadan dan tidak multitafsir.
- b. Terhadap butir soal yang sangat rendah daya pembedanya, peneliti melakukan hal sebagai berikut:
 - 1) Soal dikaji ulang secara cermat dan diperbaiki serta dianalisis, akankah soal itu mungkin memiliki daya beda yang semakin membaik, jika memungkinkan,bisa dipakai lagi..
 - 2) Soal disimpan, direnungkan kembali atau sekalian dibuang
- c. Perlu penelitian serupa seperti analisis butir soal *mahaarah qiroo'ah dan mahaaroh kitaabah* pada TOAFL UIN Maliki Malang.