

Received : 22-02-2022	Accepted : 14-06-2022
Published : 30-06-2022	Doi : 10.32699/liar.v6i1.2549

## Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal TOAFL Universitas Hasyim Asy'ari Tebuireng Jombang

Lailatul Qomariyah

Universitas Hasyim Asy'ari Tebuireng Jombang, Indonesia

Corresponding E-Mail: lailatulqomariyah.unhasy@gmail.com

### Abstract

*Testing is one of the ways to evaluate students' achievements and mastery of their skills. Through this test, the teacher will know how to measure their achievement and identify their strengths and weaknesses. Therefore, the items of test questions must be of a good quality, so that it can know the extent to which the desired educational goals have been achieved. This study is to describe the quality of test of Arabic as a foreign language UNHASY in terms of difficulty index and discriminating power. The design of this research is descriptive quantitative. The method that is used in collecting data is documentation method. The result of this study indicated that, 1) TOAFL UNHASY items have a quality level of difficulty as many as 22 items already on the medium difficulty index; 2) TOAFL UNHASY items have a discriminating power that is qualified for a total of 44 items are in distinguishing adequate. This research provides suggestions, a good test instrument before being used should be discussed in a Focus Group Discussion (FGD) as well as socialization with education actors and experts in the field of measurement and evaluation.*

**Keywords:** Evaluation, Level of difficulty; discriminating power; test item; TOAFL

## A. Pendahuluan

Keberhasilan pembelajaran Bahasa Arab dipengaruhi oleh beberapa faktor, yakni guru, murid dan tujuan pembelajaran. Arifudin 2015 menjelaskan bahwa guru berperan penting dalam penentuan kualitas peserta didik baik akademisnya, kematangan emosionalnya, serta ketercapaian tujuan pembelajaran yang telah ditetapkan oleh guru. Untuk mewujudkannya, guru harus memiliki beberapa kompetensi, salah satunya yakni evaluator. Hal ini sesuai dengan Undang Undang No. 14 Tahun 2005 pasal 10 ayat 1 bahwa guru sebagai penentu kesuksesan proses belajar mengajar harus berkompeten dalam merancang, melaksanakan, mengevaluasi, dan merevisi butir soal yang digunakan dalam kegiatan evaluasi yang merupakan penentu keberhasilan atau tidak berhasil proses pembelajaran yang telah dilakukan. Hal senada disampaikan oleh Mutholib, 2013, bahwa kegiatan ini adalah bagian dari beberapa komponen pembelajaran yang saling terkait yakni perencanaan pembelajaran, pelaksanaannya dan penilaiannya. Jadi dengan kegiatan evaluasi akan diketahui sejauhmana ketercapaian siswa dalam tujuan pembelajaran yang telah ditetapkan. Ainin, 2016, menambahkan bahwa kegiatan ini juga sebagai pedoman dalam memberikan umpan balik (*feed back*) untuk memperbaiki sistem pembelajaran.

Dengan kompetensi tersebut, guru harus memilih alat instrumen yang sesuai dalam kegiatan penilaian dengan menyesuaikan tujuan yang akan dicapai, lalu menentukan tujuan pembelajaran yang akan dinilai, kemudian menentukan kisi-kisi soalnya dan menyusun butir soalnya, lalu dilanjutkan dengan kegiatan penilaian dan hasilnya akan didokumentasikan untuk mengetahui perkembangan siswa dan akan menjadi pedoman bagi guru untuk menentukan proses pembelajaran selanjutnya. Hal ini sesuai dengan pendapat Arikunto, 2013 dan Solichin, 2017 bahwa kegiatan evaluasi dapat menggunakan alat atau instrumen baik berupa tes ataupun non tes yang digunakan untuk mengukur sejauhmana tujuan pembelajaran telah dicapai dan menjadikan pelaksanaan penilaian lebih terarah menuju pada tujuan pembelajaran.

Tes (اختبار) adalah salah satu instrument pengukur tingkatan kemampuan peserta didik. Tes ini berisi sejumlah butir soal yang disusun oleh guru secara tepat sehingga mampu menguji kemampuan siswa yang ingin dicapai dan mewakili seluruh tujuan pembelajaran yang telah ditetapkan (Purwanti, 2014). Hal ini sesuai pendapat Kadek (Primayana, Kadek Hengki, 2020) bahwa tes ini disusun guru untuk mengetahui efektivitas dan efisiensi aktifitas belajar mengajar yang akan menjadi pijakan guru menentukan keberhasilan pembelajaran dan mencapai tujuan pembelajaran yang di tetapkan. Sehingga sebagai dasar dalam kegiatan memperbaiki dan mengembangkan keberlanjutan proses pembelajaran.

Arifin, 2012, Arikunto, 2013 dan Munthe, 2015 dan Ma'arif berpendapat sama dengan pandangan sebelumnya bahwa tes (اختبار) sebagai alat untuk mengukur dan menilai yang lebih sering digunakan dalam kegiatan pembelajaran. Kegiatan ini memiliki tujuan untuk memperoleh data sejauhmana siswa mencapai hasil belajarnya dan perkembangan belajar siswa.

TOAFL (*Test of Arabic Foreign Language*/ اختبارات الكفاية اللغوية لغير العرب) merupakan tes standar yang diujikan kepada mahasiswa UNHASY sebagai alat pengukur kemampuan mereka dalam Berbahasa Arab. Tes ini diwajibkan sebagai prayarat kelulusan mahasiswa program strata 1 di UNHASY. Tes ini dikembangkan oleh UPB Unhasy terdiri dari 164 soal yang berbentuk 160 pilihan ganda dan 4 uraian. Soal pilihan ganda terdiri dari 40 soal untuk mengukur kemampuan menyimak, 50 untuk mengukur kemampuan membaca, 40 soal untuk mengukur kemampuan qoidah Bahasa Arab, 30 soal untuk mengukur perbendaharaan mufrodat dan 4 soal untuk mengukur penguasaan dalam mengarang bahasa Arab.

Pada awalnya, butir soal TOAFL disusun oleh Tim Dosen Bahasa Arab yang telah digunakan sejak tahun 2014 untuk mengetahui sejauhmana ketercapaian mahasiswa dalam kemampuan berbahasa Arab setelah dilakukannya program intensif bahasa arab mulai jam 07.00-08.00 WIB dengan bobot 1 sks dalam waktu dua hari selama satu tahun. Dengan

adanya standarisasi tes ini diharapkan mampu menekan mahasiswa untuk memiliki ketrampilan berbahasa yang matang.

Pada perkembangannya, tes ini juga diujikan bagi para dosen, yang mana sertifikat yang diperoleh dapat dipergunakan sebagai prasyarat peningkatan kualifikasi dosen yakni untuk prasyarat mendapatkan sertifikat pendidik dan kualifikasi akademiknya yakni mengikuti program doktoral. Hasil program tersebut masih belum begitu tampak, hal ini dapat terlihat masih belum ada standarisasi kemampuan mahasiswa dalam bidang bahasa Arab pada semua prodi di bawah naungan UNHAS yang diketahui dengan adanya perbedaan penggunaan buku rujukan pada materi Bahasa Arab 1 dan Bahasa Arab 2.

Oleh sebab itu perlulah dilakukan telaah kembali soal TOAFL untuk mengungkap dan mempertegas kualitas butir soalnya. Fenomena ini menimbulkan sebuah pertanyaan sudahkah butir soal TOAFL yang telah disusun oleh Tim dosen Bahasa Arab memenuhi alat tes yang berkualitas sehingga dapat sebagai pengukur kemampuan siswa. Bisa jadi mahasiswa belum mencapai skor minimal TOAFL bukan karena mereka belum memahami materi Bahasa Arab yang telah dipelajari akan tetapi karena butir soal belum memenuhi kriteria berkualitas. Sehingga masih belum bisa mengukur sejauhmana hasil belajar mahasiswa. Oleh sebab itu, perlulah dianalisis kelayakan butir soal yang telah dikembangkan baik tingkat kesukaran dan daya pembedanya. Untuk itu, dengan dilakukan kegiatan ini maka akan diketahui kualitas tiap butir soal pada tes yang telah dilakukan sehingga memberikan informasi bagi dosen baik dalam menyusunnya, mengembangkannya, dan merevisi butir soal tes yang belum valid dan mempertahankan kualitas butir soal yang baik. Dengan analisis kualitas butir soal, maka akan diidentifikasi kekurangan-kekurangan butir soal (بنود الاختبار) tersusun sehingga menjadi pedoman dalam penyusunan meningkatkan kualitas tes.

Hasil analisa soal TOAFL tersebut diharapkan juga memberikan masukan berharga bagi lembaga bahasa yang menjadi umpan balik guna

memperbaiki, membenahi, dan menyempurnakan kembali pada butir soal TOAFL yang disusun sehingga alat tes tersebut betul betul akan berfungsi sebagai alat pengukur ketrampilan berbahasa yang berkualitas baik dan terstandarisasi. Mustofa (2012) juga memberikan penjelasan tentang fungsi dari analisis butir soal dengan lebih terperinci, yakni untuk 1) mengetahui soal yang sulit untuk dikerjakan peserta tes, sehingga guru akan membimbing secara khusus pada materi yang dianggap sulit; 2) menganalisa tingkat daya beda soal sehingga sebagai landasan guru dalam penentuan soal mana yang perlu dipertahankan dan perlu direvisi.

Pendapat senada dikemukakan oleh Fitriawanati, 2015 bahwa kegiatan analisis kualitas soal dilakukan untuk mendapatkan informasi ketercapaian pembelajaran yang menjadi pedoman bagi guru untuk menentukan langkah selanjutnya yakni mempertahankan butir soal yang telah disusun, memperbaikinya, membenahinya dan menyempurnakannya sehingga butir soal tersebut dapat digunakan sebagai instrumen penilaian yang layak digunakan pada kegiatan pembelajaran selanjutnya. Hal ini juga senada dengan pendapat Surapranata, 2009 dan Daryanto, 2008 bahwa dalam kegiatan evaluasi perlu adanya kegiatan analisa butir soal yang diujikan agar diketahui sejauhmana materi dapat tersampaikan kepada siswa.

Berdasarkan realita ini, peneliti ingin menganalisis kualitas butir soal pada instrumen tes TOAFL yang telah disusun dan digunakan sebagai tes Standarisasi kemampuan mahasiswa Universitas Hasyim Asy'ari dalam bidang Bahasa Arab.

Prinsip kebaruan mendasari penelitian analisis kualitar butir soal TOAFL ini adalah hasilnya ini bernilai tepat guna yakni dengan adanya revisi butir soal yang tidak layak dengan menghapus dan memperbaikinya dan butir soal yang layak tetap dipertahankan. Hasil revisi tes nanti akan diajukan HKI nya sehingga bisa digunakan untuk tes standarisasi kemampuan bahasa Arab bagi mahasiswa dan dosen di lingkungan UNHASY.

Penelitian ini dilaksanakan untuk menganalisa kualitas butir soal tes TOAFL Bahasa Arab yang telah disusun oleh Tim pengembang Bahasa Arab di UPB Universitas Hasyim Asy'ari Tebuireng Jombang yang berfokus pada kemahiran membaca (قراءة) dari segi kuantitatifnya yakni tingkat kesukaran dan daya pembedanya yang berjumlah 50 butir soal. Tes ini diperlukan oleh mahasiswa UNHASY untuk tujuan akademik yakni sebagai prasyarat kegiatan PPL, KKN, Skripsi dan Beasiswa. Penelitian sejenis tentang “Analisis Butir Soal Toafl (*Test Of Arabic Foreign Language*)” telah banyak ditemukan antara lain penelitian Fatimah Depi Susanty, 2013 dan Harahap, 2018.

Fatimah Depi Susanti dalam penelitiannya menjelaskan bahwa soal tes hasil belajar yang disusun mayoritas berbentuk essay dan objektif yang disusun dilengkapi dengan kisi-kisi soal yang sesuai dengan tujuan pembelajaran. Tes yang berbentuk essay divalidasi sepenuhnya kualitasnya oleh dosen sedangkan tes yang berbentuk objektif ditelaah kualitasnya hanya oleh sebagian kecil dosen.

Partomuan Harahap dalam penelitiannya diperoleh hasil bahwa soal TOAFL IAIN Curup yang disusun memiliki prosentase tingkat kesukaran (معامل الصعوبة) berkategori baik karena lebih dari 50% indeks memenuhi standar, dan prosentase validitas soal tersebut menunjukkan lebih 50% soal berkualitas valid. Sedangkan reliabilitas menunjukkan bahwa soal TOAFL memiliki kategori reliabilitas tinggi yakni 0,87.

Kajian penelitian ini berbeda dengan beberapa penelitian di atas. Penelitian ini perlu dilakukan untuk memperoleh hasil tentang informasi kualitas soal TOAFL UNHASY yang diharapkan akan membantu untuk mengidentifikasi kualitas butir soal baik per itemnya atau secara keseluruhan sehingga dijadikan pedoman dalam perbaikan butir soal dan mengonstruksi program pembelajaran secara kuantitatif.

## B. Metode

Pendekatan penelitian adalah penelitian kuantitatif. Peneliti ingin mengetahui kualitas butir soal TOAFL yang disusun oleh tim Dosen Bahasa Arab universitas Hasyim Asy'ari untuk mengukur kemampuan Bahasa Arab Mahasiswa UNHASY. Data dalam penelitian ini diambil melalui teknik dokumentasi yakni berupa dokumen naskah hasil TOAFL tahun 2019/2020 yang diujikan kepada 221 mahasiswa yang menjadi sampel dalam penelitian ini dan diambil dengan teknik simple random sampling, lembar soal TOAFL Bahasa Arab; dan data kurikulum Bahasa Arab.

Analisis kuantitatif dilakukan untuk mengetahui dan menganalisis tingkat kesukaran (معامل الصعوبة) dengan menggunakan proporsi menjawab benar (*proportion correct*) dan kemudian ditentukan rentangan indeks daya beda (معامل التمييز). Analisa Tingkat kesukaran menggunakan rumus berikut,

$$P. (\text{معامل الصعوبة}) = \frac{\sum B}{N}$$

Ket

P : Tingkat kesukaran

$\sum B$  : Jumlah peserta tes menjawab benar

N : Jumlah peserta tes

Kemudian untuk melakukan penafsiran tingkat kesukaran dikonversi pada patokan berikut,

$P > 0,70$  : Mudah (سهل)

$0,30 \leq p \leq 0,70$  : Sedang (معتدل)

$P < 0,70$  : Sukar (صعب)

Adapun untuk mengukur daya beda, peneliti menggunakan rumus berikut,

$$D (\text{معامل التمييز}) = \frac{\sum \text{Benar Upper} - \sum \text{Benar lower}}{\sum \text{Kelompok}}$$

Kemudian untuk menentukan rentangan indeks daya beda suatu tes dapat merujuk rumus yang dikemukakan Djiwandono (dalam Siti, 2019: 89) sebagai berikut,

(ذات تمييز عال) Baik :	$0,50 < D$ (معامل التمييز)
(ذات تمييز جيد) Sedang :	$0,50 \leq D$ (معامل التمييز)
(ضعيفة التمييز) Kurang :	$0,20 < D$ (معامل التمييز)
(السالب/غير مقبول للتمييز) Negatif :	0

### C. Hasil dan Pembahasan

#### Tingkat Kesukaran Butir Soal TOAFL UNHASY

Kesukaran (معامل الصعوبة) butir soal dianalisis agar diketahui kualitas butir soal yang dinyatakan baik. Butir soal baik akan diketahui dari hasil kemampuan mahasiswa menjawab tiap butir soalnya yang menunjukkan bahwa soal itu seimbang yakni butir soal tidak terlalu mudah untuk dijawab dan juga tidak terlalu sukar. Ni Kadek Ratna Wati, I Nengah Suandi, 2015 dan irfan (2008, العرفان) memiliki pendapat yang sama bahwa soal dikatakan baik jika butir soal yang disusun seimbang sehingga siswa tetap semangat dan fokus dalam menyelesaikan soal karena terhindar dari rasa putus asa ketika menghadapi soal yang terlalu sulit.

Peneliti akan menganalisa tingkat kesukaran butir soal dengan membaginya sesuai dengan ketrampilan yang diukur dengan menggunakan proporsi menjawab benar. Berikut akan dianalisa tingkat kesukaran butir soal TOAFL UNHASY yang berfokus pada soal ketrampilan membaca yang terdiri dari 50 butir soal yang diberikan kepada 221 mahasiswa. Kemudian berdasarkan hasil jawaban benar dan salah diketahui tingkat kesukaran dari tiap butir soal yang kemudian hasilnya ditafsirkan kriteria kesukaran. Berikut hasil penafsiran tingkat kesukaran pada tiap butir soal.

Tabel 1. Penafsiran hasil Tingkat Kesukaran

No	Tingkat Kesukaran (معامل الصعوبة)	Penafsiran
1	0.624	Sedang (معتدل)
2	0.814	Mudah(سهل)
3	0.643	Sedang (معتدل)
4	0.733	Mudah(سهل)
5	0.656	Sedang (معتدل)
6	1	Mudah(سهل)
7	0.660	Sedang (معتدل)
8	0.542	Sedang (معتدل)
9	0.217	Sukar(صعب)
10	0.733	Mudah(سهل)
11	0.859	Mudah(سهل)
12	0.932	Mudah(سهل)
13	0.253	Sukar(صعب)
14	0.733	Mudah(سهل)
15	0.733	Mudah(سهل)
16	0.868	Mudah(سهل)
17	1	Mudah(سهل)
18	0.868	Mudah(سهل)
19	0.868	Mudah(سهل)
20	0.497	Sedang (معتدل)
21	0.660	Sedang (معتدل)
22	0.868	Mudah(سهل)
23	0.733	Mudah(سهل)
24	0.733	Mudah(سهل)
25	0.868	Mudah(سهل)
26	0.868	Mudah(سهل)
27	0.886	Mudah(سهل)
28	0.886	Mudah(سهل)
29	0.873	Mudah(سهل)
30	0.542	Sedang (معتدل)
31	0.814	Mudah(سهل)
32	0.660	Sedang (معتدل)

33	0.308	Sedang (معتدل)
34	0.475	Sedang (معتدل)
35	0.230	Sukar(صعب)
36	0.733	Mudah(سهل)
37	0.547	Sedang (معتدل)
38	0.488	Sedang (معتدل)
39	0.660	Sedang (معتدل)
40	0.660	Sedang (معتدل)
41	0.660	Sedang (معتدل)
42	0.0669	Sedang (معتدل)
43	0.325	Sedang (معتدل)
44	0.733	Mudah(سهل)
45	0.533	Sedang (معتدل)
46	0.325	Sedang (معتدل)
47	0.271	Sukar(صعب)
48	0.266	Sukar(صعب)
49	0.660	Sedang (معتدل)
50	0.651	Sedang (معتدل)

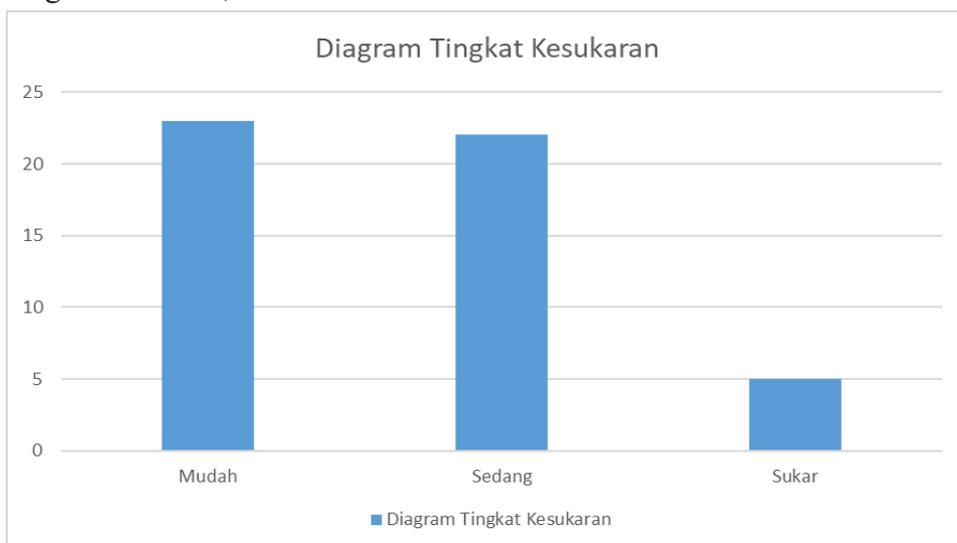
Berdasarkan tabel 1 penafsiran di atas diketahui hasil analisa tingkat kesulitan pada butir soal Ketrampilan membaca yang dapat dilihat pada tabel berikut.

Tabel 2. Hasil Analisis Soal Ketrampilan Membaca

	Kategori	Jumlah	Keterangan/Nomor soal
Analisa Tingkat Kesulitan	Mudah (سهل)	23	2,4,6,10,11,12,14,15,16,17,18, 19,22,23,24,25,26,27,28,29,31, 36,44
	Sedang (معتدل)	22	1,2,,7,8,20,21,30,32,33,34,37,38 ,39,40,41,42,43,45,46,49,50
	Sukar (صعب)	5	9,13,35,47,48

Tabel 2. Menunjukkan bahwa diantara soal no 1-50 terdapat 23 berkategori mudah(سهل) , 22 berkategori sedang (معتدل) , dan 5 butir

soal berkategori sukar(صعب). Untuk lebih mudahnya dapat dilihat pada diagram berikut,



Berdasarkan diagram di atas diketahui bahwa prosentasi proporsi tingkat kesukaran soal TOAFL UNHASY pada ketrampilan membaca berkategori cukup karena 40%-50% butir soal yang memiliki tingkat kesukaran sedang.

Soal TAOFL ini dikembangkan dengan memiliki tingkat kesukaran sedang disesuaikan dengan tujuan diperuntukkannya tes diberikan yakni bagi Mahasiswa Semester 2 dari Semua Program Studi di Universitas Hasyim Asy'ari yang mana Mayoritas dari mahasiswa adalah bukan dari program Keagamaan. Jadi diadakannya tes ini adalah untuk mengetahui standar kemampuan Bahasa Arab Mahasiswa yang akan menjadi panduan bagi setiap prodinya untuk melakukan pengembangan kemampuan Bahasa Arab Mahasiswa yang akan diintegrasikan dalam Mata Kuliahnya yakni Mata kuliah Bahasa Arab Dasar dan Bahasa Arab lanjut.

Hal ini sesuai dengan pendapat Abdul Mutholib (Mutholib, 2013) bahwa dalam menyusun soal Bahasa Arab agar memenuhi tes standar maka harus memiliki tingkat kesukaran sedang karena menyesuaikan dengan materi yang diajarkan.

### Daya Pembeda (معامل التمييز) Butir Soal TOAFL UNHASY

Pramono & Kusriani, 2018 dan Al Bahrum menjelaskan bahwa daya pembeda (معامل التمييز) soal merupakan kemampuan butir soal yang disusun untuk menunjukkan perbedaan antara mahasiswa yang pandai (kelompok berkemampuan tinggi/المجموعة العليا) dengan mahasiswa yang kurang pandai (kelompok berkemampuan rendah/المجموعة الدنيا). Pada penelitian ini, peneliti menentukan daya beda (معامل التمييز) dengan mengambil 27 % menjadi kelompok atas (المجموعة العليا) dan 27 % untuk kelompok bawah (المجموعة الدنيا) dari 221 mahasiswa yang menjadi teste. Jadi ada 60 orang dari kelompok atas dan 60 orang dari kelompok bawah. Kemudian hasil jawaban siswa dikategorikan menjadi dua yakni kelompok bawah (المجموعة الدنيا) dan atas (المجموعة العليا). Setelah data dikategorikan, Maka dilakukan analisis untuk menghitung tingkat daya beda pada tiap butir soalnya dan ditafsirkan. Berikut dijelaskan tafsiran hasil analisis tingkat daya beda.

Tabel 3. Tafsiran Tingkat Daya Beda

No	Tingkat Daya Beda	Penafsiran
1	0,5	Baik (ذات تمييز عال)
2	0,5	Baik (ذات تمييز عال)
3	0,5	Baik (ذات تمييز عال)
4	0,25	Sedang Baik (ذات تمييز جيد)
5	1	Baik (ذات تمييز عال)
6	0	Tidak ada deskriminasi
7	0,25	Sedang (ذات تمييز جيد)
8	0,25	Sedang (ذات تمييز جيد)
9	0	Tidak ada deskriminasi
10	0,25	Sedang (ذات تمييز جيد)
11	0,5	Baik (ذات تمييز عال)
12	0,5	Baik (ذات تمييز عال)
13	0,5	Baik (ذات تمييز عال)
14	0,5	Baik (ذات تمييز عال)
15	0,5	Baik (ذات تمييز عال)
16	0,5	Baik (ذات تمييز عال)

17	0	Tidak ada deskriminasi
18	0,5	Baik (ذات تمییز عال)
19	0,5	Baik (ذات تمییز عال)
20	0,5	Baik (ذات تمییز عال)
21	0,75	Baik (ذات تمییز عال)
22	0,25	Sedang (ذات تمییز جيد)
23	0,5	Baik (ذات تمییز عال)
24	0,25	Sedang (ذات تمییز جيد)
25	0,75	Baik (ذات تمییز عال)
26	0,75	Baik (ذات تمییز عال)
27	0,5	Baik (ذات تمییز عال)
28	0,5	Baik (ذات تمییز عال)
29	0,75	Baik (ذات تمییز عال)
30	0,25	Sedang (ذات تمییز جيد)
31	0,5	Baik (ذات تمییز عال)
32	0,75	Baik (ذات تمییز عال)
33	0,5	Baik (ذات تمییز عال)
34	0	Tidak ada deskriminasi
35	0,25	Sedang (ذات تمییز جيد)
36	0,75	Baik (ذات تمییز عال)
37	0,75	Baik (ذات تمییز عال)
38	0,25	Sedang (ذات تمییز جيد)
39	0,25	Sedang (ذات تمییز جيد)
40	0,5	Baik (ذات تمییز عال)
41	0,75	Baik (ذات تمییز عال)
42	-0,25	Negative (السالب)
43	0	Tidak ada deskriminasi
44	1	Baik (ذات تمییز عال)
45	0,25	Sedang (ذات تمییز جيد)
46	0,75	Baik (ذات تمییز عال)
47	0,5	Baik (ذات تمییز عال)
48	0,5	Baik (ذات تمییز عال)
49	0,5	Baik (ذات تمییز عال)
50	0,5	Baik (ذات تمییز عال)

Berdasarkan tabel di atas diketahui kategori tingkat daya beda pada tiap butir soalnya.

Tabel 4. Hasil Analisis Daya Pembeda Soal

Analisa Tingkat Kesulitan	Kategori	Jumlah	Keterangan/Nomor soal
	Baik	33	1,2,3,5,11,12,13,14,15,16,18,19,20,21,23,25,26,27,28,29,31,32,33,36,37,40,41,44,46,47,48,49,50
	Sedang	11	4,7,8,10,22,24,30,35,38,39,45
	Tidak ada Deskriminasi	5	6,9,17, 34,43
	Negatif	1	42

Berdasarkan tabel 4. Maka diketahui bahwa ada 33 soal yang berkategori baik dan 11 soal berkategori sedang. Butir soal dengan dua kategori tersebut masih dipertahankan. Butir soal yang berkategori tidak ada deskriminasi tidak mampu membedakan mahasiswa kelompok bawah dan atas yang terdiri dari 5 butir soal dan 1 soal berkategori negatif. Butir soal ini harus dibuang atau tidak dipergunakan kemudian diganti dengan soal yang lain. Berikut disajikan digram daya pembeda soal agar lebih mudah dipahami tentang kategori daya pembeda soal TOAFL yang telah dianalisis.

Berdasarkan diagram di atas diketahui bahwa prosentasi proporsi daya pembeda soal TOAFL UNHASI pada ketrampilan membaca berkategori baik karena lebih dari 50% butir soal yang memiliki daya pembeda baik (ذات تمييز عال). Hal ini menunjukkan bahwa soal yang disusun telah mampu melakukan perbedaan siswa yang menjadi testee dengan memiliki prestasi tinggi dan kelompok testee yang memiliki prestasi rendah dan juga diketahui beberapa soal yang perlu direvisi sehingga tersusun soal yang mampu memiliki kualitas yang baik.

Hal ini sesuai dengan pendapat Abdul (Mutholib, 2013), dan Solichin (Solichin, 2017) bahwa soal yang memiliki kualitas yang baik perlu memiliki daya beda yang baik yang mampu membedakan kelompok siswa yang berprestasi tinggi dan berprestasi rendah.

#### **D. Kesimpulan**

Berdasarkan analisa butir soal TOAFL UNHASY Tebuireng pada materi Ketrampilan Membaca yang telah dilakukan peneliti, maka dapat ditarik kesimpulan berikut, 1) Tingkat kesukaran soal TOAFL UNHASY dikatakan cukup baik, karena memiliki proporsi yang belum seimbang yakni, 23 soal berkategori mudah(22 , (سهل) berkategori sedang(معتدل), dan 5 butir soal berkategori sukar(صعب). Oleh sebab itu, 23 butir soal yang berkategori mudah dimungkinkan untuk dibuang atau direvisi lagi sehingga bisa dipergunakan pada kegiatan evaluasi selanjutnya 22 butir soal yang berada pada tingkat sedang bisa dipertahankan dan 5 butir soal yang berkategori sukar perlu diganti dan direvisi dengan soal yang baru yang lebih mudah. 2) Daya pembeda soal TOAFL dikatakan baik karena lebih dari 50% soalnya mampu membedakan mahasiswa kelompok atas dan bawah yakni dengan 44 soal yang perlu dipertahankan dan 6 butir soal direvisi. Penelitian ini memberikan saran, bahwa instrumen tes yang baik sebelum digunakan sebaiknya didiskusikan dalam *Focus Group Discussion* (FGD) serta sosialisasi dengan pelaku pendidikan dan pakar di bidang pengukuran dan evaluasi sehingga akan diperoleh butir soal yang berkualitas baik kualitas dan kuantitas. Hasil penelitian ini merekomendasikan bahwa diperlukan penelitian selanjutnya yakni penelitian analisis kualitas butir soal dari segi kualitatif baik dari segi materi, konstruksi dan bahasa dari butir soal yang disusun.

## Daftar Pustaka

- Ainin, M. (2016). Kesahihan Dalam Penyusunan Tes Bahasa Arab Di Madrasah/Sekolah. *Prosiding Konfererensi Nasional Bahasa Arab*, 1(2). <http://prosiding.arab-um.com/index.php/konasbara/article/view/75>
- Arifin, Z. (2012). *Evaluasi Pembelajaran* (2nd ed.). Direktorat Jenderal Pendidikan Islam, Kementerian Agama RI.
- Arifudin, I. S. (2015). Peranan Guru Terhadap Pendidikan Karakter Siswa Di Kelas V Sdn 1 Siluman. *PEDADIDAKTIKA: Jurnal Ilmiah Pendidikan Guru Sekolah Dasar*, 2(2), 175–186.
- Arikunto, S. (2013). *Dasar-dasar Evaluasi Pendidikan*. Bumi Aksara.
- Daryanto. (2008). *Evaluasi Pendidikan*. Rineka Cipta.
- Fitrianawati, M. (2015). Peran analisis butir soal guna meningkatkan kualitas butir soal, kompetensi guru dan hasil belajar peserta didik. *Prosiding Seminar Nasional Pendidikan PGSD UMS & HDPGSDI Wilayah Jawa*, 5(3), 282-295. ISBN 978-602-70471-2-9 PERAN. <http://hdl.handle.net/11617/9117>
- Harahap, P. (2018). Analisis Soal TOAFL IAIN CURUP. *Ihya' AL-Arabiyyah*, 2, 18–33.
- Ma'arif, A. S. (n.d.). Shiyagoh Ikhtibarot al Arabiyyah fi Dhou al Ikhtibar Al Muttaqin(TOAFL). *Al Bayan*, 9(2), 160–173.
- Munthe, A. P. (2015). Pentingnya Evaluasi Program di Institusi Pendidikan: Sebuah Pengantar, Pengertian, Tujuan dan Manfaat. *Scholaria : Jurnal Pendidikan Dan Kebudayaan*, 5(2), 1. <https://doi.org/10.24246/j.scholaria.2015.v5.i2.p1-14>
- Mustofa, S. (n.d.). Analisis Butir Tes Soal TOEFL Universitas Islam Negeri Maulana Malik Ibrahim Malang. In A. M. A. B. Siti Nurul Azkiyah, Teguh Khaerudin (Ed.), *Language Teaching and Assesment for The 21 Century: Current and Trends* (pp. 131–140). Pusat Pengembangan Bahasa UIN Jakarta.

- Mutholib, A. (2013). “Analisis Butir Soal Bahasa Arab Ujian Akhir Madrasah Berstandar Nasional (UAMBN) Bagi Siswa Madrasah Ibtidaiyah (MI) Tahun Pelajaran 2012-2013.” *Jurnal Arabia*, 5(2).
- Ni Kadek Ratna Wati, I Nengah Suandi, I. W. W. (2015). Kualitas Butir Soal Ulangan Akhir Semester Ganjil Bahasa Indonesia Kelas Xi Sma Negeri 2 Singaraja Tahun Pelajaran 2015/2016 Dari Segi Taraf Kesukaran, Daya Pembeda, dan Fungsi Pengecoh. *E-Journal PBSI*, 3(2), 1–11. <https://ejournal.undiksha.ac.id/index.php/JJPBS/article/view/7202>
- Pramono, A., & Kusri, K. (2018). Analisis Butir Soal Pilihan Ganda Menggunakan Fuzzy Berdasar Data Learning Management System Studi Kasus: SMK Negeri 2 Kediri. *Cahaya Tech*, 7(1), 36. <https://doi.org/10.47047/ct.v7i1.3>
- Primayana, Kadek Hengki, I. P. S. (2020). Peran Desain Evaluasi Pembelajaran. *Widyacarya*, 4(2), 88–100. <http://jurnal.stahnmpukuturan.ac.id/index.php/widyacarya/article/view/796>
- Purwanti, M. (2014). Analisis Butir Soal Ujian Akhir Mata Pelajaran Akuntansi Keuangan Menggunakan Microsoft Office Excel 2010. *Jurnal Pendidikan Akuntansi Indonesia*, 12(2). <https://doi.org/10.21831/jpai.v12i2.2710>
- Solichin, M. (2017). Analisis Daya Beda Soal Taraf Kesukaran, Butir Tes, Validitas Butir Tes, Interpretasi Hasil Tes Validitas Ramalan dalam Evaluasi Pendidikan. *Jurnal Manajemen Dan Pendidikan Islam* 2, 2(2), 192–213.
- Surapranata. (2009). *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes*. Remaja Rosdakarya.
- Susanty, F. D. (2013). Analisis Validasi Soal Tes Hasil Belajar pada Pelaksanaan Pembelajaran Bahasa Arab di Pusat Pengembangan Bahasa (P3B) UIN SUSKA Riau. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699.

العرفان, ز. (٢٠٠٨). الاختبار النهائي الوطني لمادة اللغة العربية بالمدرسة الثانوية  
للسنة الدراسية ٢٠٠٧-٢٠٠٨. جامعة مولانا مالك إبراهيم الإسلامية الحكومية  
مالانج.