

Analisis Butir Tes Pilihan Ganda Berbantuan *Software* Anates Mata Kuliah Menyimak Mahasiswa PBSI Universitas Riau

Anggita Putri Mulyani¹ Elvrin Septyanti² Tria Putri Mustika³

Program Studi Pendidikan Bahasa dan Sastra Indonesia, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Riau, Kota Pekanbaru, Provinsi Riau, Indonesia^{1,2,3}

Email: anggita.putri0589@student.unri.ac.id¹

Abstrak

Analisis soal tes dilakukan untuk mengetahui apakah suatu tes dapat dikatakan berkualitas baik atau buruk dengan melihat tiga hal, yaitu tingkat kesukaran, daya pembeda, keefektifan alternatif jawaban setiap soal, dan keefektifitasan distraktor hal ini untuk membedakan antara kelompok unggul dan kelompok asor. Penelitian ini dilaksanakan di Fakultas Keguruan dan Ilmu Pendidikan Universitas Riau khususnya pada Program Studi Pendidikan Bahasa dan Sastra Indonesia pada mata kuliah menyimak. Peneliti menggunakan metode analisis kuantitatif yang dilakukan dengan bantuan program Anates versi 4.0.2 yang hasilnya meliputi reliabilitas, daya pembeda, tingkat kesukaran soal, dan fungsi pengecoh dengan populasi seluruh mahasiswa angkatan 2024 yang berjumlah 119 dan sampel sebanyak 84 siswa yang diambil secara random sampling. Teknik pengumpulan data menggunakan tes dan dokumentasi. Melalui rekapitulasi analisis butir soal, reliabilitas tes sebesar 0,37 berada pada kategori rendah yang berarti tes tersebut tidak dapat digunakan berulang kali pada responden yang sama dengan hasil yang sama. Daya pembeda soal berada pada kategori sedang dengan persentase 57%. Tingkat kesukaran tes didominasi kategori sedang dengan skor 43%. Efektivitas pengecoh pada soal tes sebesar 64% berkategori dapat diterima. Penelitian ini menemukan bahwa soal tes pilihan ganda pada mata pelajaran menyimak perlu ditingkatkan dan diperbaiki terutama pada soal yang tidak relevan dan terlalu sulit, daya pembeda yang lemah, dan pengecoh yang tidak efektif. Perbaikan ini dapat membantu memastikan bahwa tes tersebut dapat diandalkan, valid, dan efektif dalam mengukur pemahaman siswa terhadap materi.

Kata Kunci: Tes, Menyimak, Anates



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

PENDAHULUAN

Pendidikan merupakan suatu hal yang fleksibel dan dapat diperoleh dimana saja dan merupakan upaya efisien untuk mengembangkan potensi dan kualitas berpikir seseorang, baik yang diperoleh melalui lingkungan sekolah maupun di luarnya. Perguruan tinggi bertujuan untuk membentuk peserta didik yang mampu berpikir aras tinggi, inovatif, dan kreatif melalui penyelenggaraan pendidikan, pelaksanaan penelitian, dan pengabdian kepada masyarakat, sebagaimana yang tercantum dalam Undang-Undang Dasar nomor 12 Tahun 2012 tentang Pendidikan Tinggi, pasal 1 ayat (1), dan ayat (9); dan ayat 4b. Peraturan Menteri Pendidikan dan Kebudayaan tentang Standar Nasional Pendidikan Tinggi menyatakan bahwa aspek interaktif, tematik, efektif, dan konseptual terhadap mahasiswa diperlukan dalam proses pembelajaran. Semua atribut tersebut perlu dimaksimalkan dalam kegiatan pendidikan, khususnya menyimak yang merupakan salah satu keterampilan berbahasa yang terdapat dalam kurikulum pendidikan tinggi.

Sebelum mempelajari kemampuan linguistik lainnya, anak secara alami sudah mahir dalam mendengarkan (Prihatin, 2017). Pada umumnya kegiatan menyimak digunakan lebih banyak daripada menulis, membaca, dan berbicara dalam aktivitas sehari-hari. Adler (Hermawan, 2012) mencatat bahwa 53% aktivitas komunikasi didominasi oleh menyimak,

berbicara 16%, membaca 17%, dan menulis 14%. (Hermawan, 2012) menyatakan bahwa dibandingkan dengan keterampilan komunikasi lainnya, mata pelajaran menyimak kurang mendapat prioritas di banyak sekolah. Dua kemampuan dasar yang sering tercakup dalam kurikulum dari taman kanak-kanak hingga sekolah menengah atas adalah membaca dan menulis. Kemampuan berbicara dan mendengarkan termasuk dalam kompetensi yang diajarkan pada jenjang pendidikan tinggi. Tenaga pendidik tentunya sangat prihatin dengan hal ini karena menyimak adalah salah satu keterampilan berbahasa yang 50% digunakan dalam kehidupan sehari-hari. Selain mengajar, melatih dan membimbing tugas lain tenaga pendidik terhadap mahasiswa adalah memberikan penilaian serta mengevaluasi. Kemampuan mengevaluasi merupakan salah satu kompetensi profesional yang perlu dimiliki seorang tenaga pendidik, baik dalam penilaian hasil pembelajaran maupun proses pembelajaran. Evaluasi menurut (Soulisa et al, 2022:2) adalah proses pengumpulan informasi yang diperlukan untuk memastikan ruang lingkup dan cara pembelajaran guna melakukan penilaian dan penyesuaian yang tepat untuk memaksimalkan hasil. Tes dan ujian identik dengan istilah "evaluasi" dalam konteks pembelajaran.

Tes adalah alat yang digunakan untuk mengukur seberapa baik siswa telah belajar dengan memberi mereka serangkaian pertanyaan atau tugas untuk diselesaikan (Halik et al, 2019). Tes pilihan ganda disebut sebagai tes objektif karena penilaiannya hanya jawaban benar yang mendapat point/skor dan jawaban salah mendapat 0. Dalam tes pilihan ganda, ada lima kemungkinan jawaban: A, B, C, D, dan E. Keuntungan tes pilihan ganda adalah mudah dinilai dan mencakup berbagai mata pelajaran. Pada tes pilihan ganda, prosedur konvensional, kompensasi, dan penalti adalah beberapa cara untuk menghitung poin atau skor. Teknik penskoran konvensional adalah teknik menentukan skor yang paling banyak digunakan pada tes pilihan ganda, teknik ini disebut juga dengan *correct score* atau *summated rating*, dimana jawaban yang benar pada tes dijumlahkan untuk menentukan skor. Analisis butir tes dilakukan untuk mengidentifikasi apakah suatu tes dapat dikatakan berkualitas baik atau berkualitas buruk dengan melihat dari tiga hal yaitu tingkat kesukaran, daya pembeda, efektivitas jawaban alternatif untuk setiap pertanyaan, dan kemampuan membedakan antara skor tinggi dan rendah. Hal ini kemudian diperkuat dengan pendapat Suharsimi (Rahmasari & Ismiyati, 2016) bahwa suatu tes dianggap berkualitas baik apabila sudah memenuhi syarat dari aspek materi, konstruksi, bahasa, serta mempunyai validitas, reliabilitas, dan daya pembeda yang tinggi sehingga tingkat kesukaran berada pada interval sedang.

METODE PENELITIAN

Penelitian dilaksanakan di Fakultas Keguruan dan Ilmu Pendidikan Universitas Riau tepatnya di Program Studi Pendidikan Bahasa dan Sastra Indonesia pada mata kuliah menyimak. Peneliti menggunakan metode analisis kuantitatif yang dilakukan dengan bantuan program Anates versi 4.0.2 yang hasilnya meliputi reliabilitas, daya pembeda, tingkat kesulitan soal, dan fungsi pengecoh. Menurut sugiyono (2017:117-118) Sampel adalah himpunan bagian dari jumlah dan sifat-sifat yang membentuk populasi, yaitu suatu wilayah *generalisasi* yang terdiri dari obyek/subyek yang mempunyai sifat tertentu yang dipilih oleh peneliti untuk dipelajari dan selanjutnya diambil kesimpulannya. Populasi pada penelitian ini adalah seluruh mahasiswa PBSI angkatan 2024 yang berjumlah 119 mahasiswa. Sampel dalam penelitian ini diambil secara random atau yang biasa dikenal dengan *random sampling*. Cara menentukan jumlah sampel dilakukan dengan melihat tabel penentuan jumlah sampel dari populasi tertentu yang dikembangkan oleh Isaac dan Michael berdasarkan tingkat kesalahannya. Populasi dalam penelitian ini berjumlah 119 orang dengan tingkat kesalahan 5%, artinya peneliti mengkehendaki kepercayaan sampel terhadap populasi adalah 95%. Berdasarkan tabel penentuan jumlah sampel dengan tingkat kesalahan 5% didapati sampel berjumlah= 84.

Table 1. Populasi dan Sampel Penelitian

Kelas	Populasi	Sampel	
PBSI 2024 A	41	41/119 x 84	28,9
PBSI 2024 B	40	40/119 x 84	28,2
PBSI 2024 C	38	38/ 119 x 84	26,8
JUMLAH	119	83,9	

Penelitian ini menggunakan indikator reliabilitas tes, daya pembeda, tingkat kesukaran tes, dan peran pengecoh, serta menggunakan kualitas ujian pilihan ganda pada mata kuliah menyimak mahasiswa PBSI Universitas Riau khusus angkatan 2024. Tes dan dokumentasi merupakan pendekatan yang digunakan dalam penelitian ini untuk memperoleh data berupa lembar jawaban tes yang sudah diisi oleh mahasiswa. Analisis data pada penelitian ini memanfaatkan program komputer Anates versi 4.0.2 yang dikembangkan oleh Drs. Karnoto, M. P.d., dan Yudi Wibisono, ST., sebagai alat bantu dalam menganalisis butir soal dengan kriteria dan aspek yang lengkap tanpa harus menggunakan alat bantu terjemahan karena program ini keseluruhannya menggunakan bahasa Indonesia.

Reliabilitas (Sugiyono, 2017:173) adalah alat untuk menghitung apakah suatu instrumen dalam hal ini tes dapat digunakan berulang kali untuk mengukur objek/responden yang sama dengan hasil yang sama pula, dengan kata lain reliabilitas merupakan suatu tolak ukur untuk sebuah instrumen dalam menentukan apakah instrumen tersebut dapat digunakan berulang kali. Dalam penelitian ini untuk menentukan reliabilitas instrumen tes pilhan ganda digunakan rumus KR. 20 (Kuder Richardson) sebagai berikut:

$$r_{11} = \left(\frac{k}{k-1} \right) \left(\frac{s^2 - \sum pq}{s^2} \right)$$

Keterangan

r_{11} : reliabilitas tes

k : jumlah item dalam instrumen

s : standar deviasi

p : jumlah subjek yang menjawab benar

q : jumlah subek yang menjawab salah

$\sum pq$: jumlah perkalian dari p dan q

Berikut tabel kategori reliabilitas untuk sebuah instrumen

Table 2. Tabel Indeks Reliabilitas Instrumen

Nilai Keandalan	Keterangan
$r_{11} < 0,20$	Sangat rendah
$0,20 < r_{11} < 0,40$	Rendah
$0,40 < r_{11} < 0,70$	Sedang
$0,70 < r_{11} < 0,90$	Tinggi
$0,90 < r_{11} < 1,00$	Sangat tinggi

Daya pembeda Arikunto (Suharsimi, 2013: 213) adalah kemampuan untuk suatu tes dalam mengelompokkan responden yang berkemampuan tinggi (unggul) dan responden berkemampuan rendah (asor). Rumus yang digunakan untuk mengetahui daya pembeda suatu tes atau soal adalah sebagai berikut:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B}$$

Keterangan

D : daya pembeda

BA : jumlah benar kelompok unggul
 BB : jumlah benar kelompok asor
 JA : jumlah peserta kelompok unggul
 JB : jumlah peserta kelompok asor

Setelah memperoleh hasil dari perhitungan rumus tersebut, untuk menentukan kategori daya pembeda digunakanlah tabel berikut:

Table 3. Tingkatan Daya Pembeda

Daya pembeda	keterangan
0-0,20	Butir soal memiliki daya pembeda lemah
0,21-0,40	Butir soal memiliki daya pembeda sedang
0,41-0,70	Butir soal memiliki daya pembeda baik
0,71-1,00	Butir soal memiliki daya pembeda sangat baik
Bertanda negatif (-)	Butir soal memiliki daya pembeda buruk

Tingkat kesulitan tes (Arikunto, 2013: 223) adalah suatu perlakuan untuk menentukan apakah suatu instrumen dalam hal ini adalah tes termasuk kedalam golongan tes yang sulit atau tidak, dengan menggunakan rumus sebagai berikut:

$$P = \frac{B}{J_s}$$

Keterangan

P : tingkat kesukaran

B : jumlah peserta yang menjawab benar

JS : jumlah peserta keseluruhan

Berikut adalah kategori tingkat kesulitan soal:

Table 4. Kategori Tingkat Kesulitan Soal

Nilai	Kategori
$0,00 < P < 0,20$	Sangat sulit
$0,21 < P < 0,40$	Sulit
$0,41 < P < 0,60$	Sedang
$0,61 < P < 0,90$	Mudah
$0,91 < P < 1,00$	Sangat mudah

Daya pengecoh/distraktor adalah opsi-opsi yang bukan kunci jawaban dan merupakan suatu syarat yang harus ada dalam instrumen berbentuk objektif atau tes. Pengecoh dianggap diterima apabila responden dengan jumlah ideal memilih pengecoh dalam tes tersebut. Berikut rumus untuk menentukan indeks daya pengecoh/distraktor:

$$D = \frac{A}{N} \times 100\%$$

Keterangan

D : tingkat distraktor

A : jumlah peserta yang memilih opsi tersebut

N : jumlah peserta keseluruhan

Berikut kategori dari indeks daya pembeda/distraktor

Table 5. Indeks Keefektifan Distraktor

Nilai	Kategori
$D > 5\%$	diterima

5% > D > 0	revisi
D = 0	ditolak

Penilaian validitas terhadap butir tes menyimak dilakukan oleh peneliti dan beberapa orang ahli (validator) untuk memvalidasikan butir tes yang ditinjau dari segi materi, konstruk, dan bahasa melalui angket penilaian yang kemudian akan dinilai dengan memberikan tanda centang (√) sesuai dengan skala penilaian yang dilampirkan dalam angket tersebut.

HASIL PENELITIAN DAN PEMBAHASAN

Table 6. Rekapitulasi Analisis Butir Tes Mata Kuliah Menyimak Mahasiswa PBSI UNRI

Rata-rata : 16,55							
Reliabilitas : 0,37							
no	Daya Pembeda		Tingkat Kesulitan		Efektifitas Pengecoh		
	angka	kategori	angka	kategori	angka	indeks (%)	kategori
1	0,22	sedang	0,25	sulit	A. 11	13%	diterima
					B. 10	12%	diterima
					C. 1	1%	revisi
					D. 21**	25%	diterima
					E. 41	49%	diterima
2	0,13	lemah	0,26	sulit	A. 49	58%	diterima
					B. 4	5%	diterima
					C. 22**	26%	diterima
					D. 6	7%	diterima
					E. 2	2%	revisi
3	0,17	lemah	0,92	sangat mudah	A. 4	5%	diterima
					B. 77**	92%	diterima
					C. 1	1%	revisi
					D. 1	1%	revisi
					E. 1	1%	revisi
4	0,17	lemah	0,24	sulit	A. 0	0	tolak
					B. 20**	24%	diterima
					C. 1	1%	revisi
					D. 61	73%	diterima
					E. 2	2%	revisi
5	0,35	sedang	0,4	sedang	A. 0	0	tolak
					B. 2	2%	revisi
					C. 34**	40%	diterima
					D. 45	54%	diterima
					E. 3	4%	revisi
6	0,26	sedang	0,55	sedang	A. 3	4%	revisi
					B. 46**	55%	diterima
					C. 10	12%	diterima
					D. 1	1%	revisi
					E. 24	29%	diterima

7	0,26	sedang	0,3	sulit	A. 25**	30%	diterima
					B. 2	2%	revisi
					C. 9	11%	diterima
					D. 47	56%	diterima
					E. 1	1%	revisi
8	0,17	lemah	0,9	sangat mudah	A. 2	2%	revisi
					B. 5	6%	diterima
					C. 76**	90%	diterima
					D. 1	1%	revisi
					E. 0	0	tolak
9	0,13	lemah	0,94	sangat mudah	A. 79**	94%	diterima
					B. 3	4%	revisi
					C. 2	2%	revisi
					D. 0	0	tolak
					E. 0	0	tolak
10	0,22	sedang	0,87	mudah	A. 3	4%	revisi
					B. 3	4%	revisi
					C. 73**	87%	diterima
					D. 1	1%	revisi
					E. 3	4%	revisi
11	0,35	sedang	0,74	mudah	A. 3	4%	revisi
					B. 14	17%	diterima
					C. 62**	74%	diterima
					D. 4	5%	diterima
					E. 1	1%	revisi
12	0	lemah	0,3	sulit	A. 1	1%	revisi
					B. 13	15%	diterima
					C. 15	18%	diterima
					D. 25**	30%	diterima
					E. 30	36%	diterima
13	0,35	sedang	0,43	sedang	A. 36**	43%	diterima
					B. 8	10%	diterima
					C. 22	26%	diterima
					D. 8	10%	diterima
					E. 10	12%	diterima
14	0,09	lemah	0,43	sedang	A. 7	8%	diterima
					B. 10	12%	diterima
					C. 36**	43%	diterima
					D. 29	35%	diterima
					E. 1	1%	revisi
15	0,22	sedang	0,63	sedang	A. 10	12%	diterima
					B. 5	6%	diterima
					C. 53**	63%	diterima
					D. 13	15%	diterima

					E. 3	4%	revisi
16	0,39	sedang	0,71	mudah	A. 15	18%	diterima
					B. 60**	71%	diterima
					C. 3	4%	revisi
					D. 5	6%	diterima
					E. 1	1%	revisi
17	0,22	sedang	0,55	sedang	A. 13	15%	diterima
					B. 5	6%	diterima
					C. 13	15%	diterima
					D. 7	8%	diterima
					E. 46**	55%	diterima
18	0,09	lemah	0,57	sedang	A. 11	13%	diterima
					B. 8	10%	diterima
					C. 6	7%	diterima
					D. 48**	57%	diterima
					E. 10	12%	diterima
19	0,22	sedang	0,77	mudah	A. 9	11%	diterima
					B. 5	6%	diterima
					C. 65**	77%	diterima
					D. 3	4%	revisi
					E. 2	2%	revisi
20	0,35	sedang	0,43	sedang	A. 2	2%	revisi
					B. 30	36%	diterima
					C. 36**	43%	diterima
					D. 2	2%	revisi
					E. 14	17%	diterima
21	0,04	lemah	0,02	sangat sulit	A. 41	49%	diterima
					B. 2**	2%	revisi
					C. 17	20%	diterima
					D. 21	25%	diterima
					E. 3	4%	revisi
22	0,3	sedang	0,52	sedang	A. 44**	52%	diterima
					B. 21	25%	diterima
					C. 2	2%	revisi
					D. 6	7%	diterima
					E. 11	13%	diterima
23	0,48	baik	0,64	sedang	A. 7	8%	diterima
					B. 5	6%	diterima
					C. 17	20%	diterima
					D. 54**	64%	diterima
					E. 1	1%	revisi
24	0,09	lemah	0,5	sedang	A. 42**	50%	diterima
					B. 17	20%	diterima
					C. 3	4%	revisi

					D. 20	24%	diterima
					E. 2	2%	revisi
25	0,39	sedang	0,27	sulit	A. 10	12%	diterima
					B. 43	51%	diterima
					C. 1	1%	revisi
					D. 23**	27%	diterima
					E. 7	8%	diterima
26	0,09	lemah	0,33	sedang	A. 28**	33%	diterima
					B. 6	7%	diterima
					C. 9	11%	diterima
					D. 3	4%	revisi
					E. 38	45%	diterima
27	0,39	sedang	0,45	sedang	A. 40	48%	diterima
					B. 0	0	tolak
					C. 38**	45%	diterima
					D. 5	6%	diterima
					E. 1	1%	revisi
28	0,17	lemah	0,93	sangat mudah	A. 2	2%	revisi
					B. 2	2%	revisi
					C. 1	1%	revisi
					D. 78**	93%	diterima
					E. 1	1%	revisi
29	0,26	sedang	0,79	mudah	A. 3	4%	revisi
					B. 5	6%	diterima
					C. 4	5%	diterima
					D. 66**	79%	diterima
					E. 6	7%	diterima
30	0,3	sedang	0,89	mudah	A. 2	2%	revisi
					B. 75**	89%	diterima
					C. 5	6%	diterima
					D. 2	2%	revisi
					E. 0	0	tolak

Berdasarkan analisis kuantitatif berbantuan program anates v4.0.2 terhadap butir tes pilihan ganda sebanyak 30 soal mata kuliah menyimak, maka validitas tes tersebut dapat dilihat melalui rekapitulasi data analisis butir tes berupa reliabilitas tes, daya pembeda, tingkat kesukaran dan keefektifan daya pengecoh/distraktor. Reliabilitas pada tes tersebut berada pada nilai 0,37 dengan rata-rata 16,55. Hal ini menunjukkan bahwa nilai keandalan untuk tes tersebut berkategori rendah dengan artian bahwa tes tersebut tidak dapat digunakan berulang kali untuk mengukur nilai objek/responden yang sama dengan hasil yang sama. (Setiyawan, 2014:345) Sebuah tes dapat dikategorikan memiliki reliabilitas yang tinggi apabila tes tersebut diberikan kepada responden yang sama dengan hasil yang tetap sama meskipun waktunya berbeda. Tinggi atau rendahnya nilai keandalan suatu tes dapat dipengaruhi oleh beberapa faktor. Menurut Crocker & Algina (Setiyawan, 2014:348) faktor yang mempengaruhi tinggi dan rendahnya reliabilitas suatu instrumen antara lain adalah tingkat kesukaran soal, panjang tes

soal, homogenitas soal dan waktu pengerjaan soal. Hal ini kemudian didukung oleh hasil penelitian Aiken et al (Setiyawan, 2014:348) dimana dalam penelitian tersebut Aiken menjelaskan bahwa faktor umum yang paling mempengaruhi tinggi rendahnya nilai reliabilitas adalah tingkat kesukaran suatu tes, yang dapat mempengaruhi jumlah variasi jawaban benar. Semakin sulit suatu tes, maka nilai reliabilitasnya akan semakin besar pula. Sebaliknya, apabila tes berkategori tidak sulit maka nilai reliabilitasnya akan rendah.

Nilai reliabilitas yang dipengaruhi oleh tingkat kesulitan tes akan berhubungan dengan level pemahaman dari responden sehingga hal ini dapat membentuk kelompok unggul dan kelompok rendah (asor). Hubungan ini bergantung pada bagaimana tes itu disajikan, apabila butir tes yang disajikan semakin banyak jumlahnya maka reabilitas semakin tinggi karena banyaknya indeks jawaban tepat. Banyaknya jumlah tes tetap harus selalu berada dalam batas pemahaman responden, jika butir tes ditemukan terlalu sulit maka akan banyak terkaan, dan dugaan-dugaan jawaban yang benar oleh responden, akan tetapi jika butir tes ditemukan terlalu mudah maka tes itu dapat menjadi tidak efektif karena dinilai tidak dapat membentuk kelompok pemahaman responden. Berdasarkan pembahasan tersebut diketahui bahwa tidak reliabelnya tes mata kuliah menyimak mahasiswa PBSI kemungkinan disebabkan oleh banyaknya jumlah soal, tingkat kesulitan dan juga kecepatan waktu pengerjaan soal. Butir tes pilihan ganda mata kuliah menyimak adalah 30 butir yang dikerjakan berdasarkan hirarki keterampilan berfikir taksonomi bloom C1 sampai dengan C5 dengan kategori HOTS. Tes pilihan ganda ditampilkan melalui program microsoft power point dengan durasi persoa/perslide adalah 30 detik, sehingga dapat memungkinkan mahasiswa tidak mencerna soal dengan baik dan hanya menerka jawaban yang dirasa paling cocok dengan pertanyaan pada masing-masing butir tes tersebut.

Secara kuantitatif dengan melihat rekapitulasi analisis butir tes tersebut dapat diketahui bahwa tingkat daya pembeda pada tes tersebut 3% berkategori baik, 57% berkategori sedang dan 40% berkategori lemah. Hal ini menjadi pertanda bahwa daya pembeda pada tes tersebut didominasi oleh kategori sedang yang berarti bahwa kualitas daya pembeda pada tes pilihan ganda mata kuliah menyimak mampu membedakan mahasiswa yang menguasai materi dengan mahasiswa yang kurang menguasai materi. Sejalan dengan hal tersebut Depdiknas (Rasmawati, 2016:328) yang menyatakan bahwa daya pembeda soal adalah kemampuan dari butir soal itu sendiri untuk menentukan kelompok unggul dan kelompok rendah. Setelah melakukan analisa terhadap tingkat kesulitan tes maka didapati 3% butir soal berkategori sangat sulit, 20% butir soal berkategori sulit, 43% butir soal berkategori sedang, 20% butir soal berkategori mudah dan 13% butir soal berkategori sangat mudah. Berdasarkan kategori yang didapatkan maka dominasi untuk tingkat kesukaran soal berada pada kategori sedang. Menurut Suharsimi (Rasmawati, 2016:328) tingkat kesulitan yang baik dalam suatu tes adalah berkategori sedang yang berarti bahwa butir tes tersebut tidak dalam tingkat yang terlalu sulit dan juga tidak dapat dikatakan terlalu mudah. Soal yang terlalu sulit dapat menyulitkan mahasiswa untuk memilih jawaban benar. Tes yang terlalu mudah juga tidak efektif karena mahasiswa akan mudah menjawab benar dan alternatif jawaban lain tidak akan dipilih. Sehingga hal ini dapat berdampak pada kualitas pengecoh/distraktor pada tes.

Tingkat efektivitas suatu pengecoh/distraktor dalam tes ini dapat dikategorikan yaitu 64% berkategori diterima, 31% berkategori diterima dengan revisi dan 5% berkategori ditolak. Dengan demikian maka pengecoh yang diterima efektif untuk digunakan pada tes-tes berikutnya, sedangkan untuk pengecoh yang berkategori revisi harus diperbaiki terlebih dahulu sebelum kemudian akan digunakan dan kategori pengecoh yang ditolak harus diganti atau dibuang sebelum digunakan. Secara keseluruhan, butir tes pilihan ganda mata kuliah menyimak perlu diperbaiki seperti soal yang tidak relevan dan terlalu sulit, daya pembeda yang

lemah karena tidak mampu mengukur beda kemampuan menguasai materi dan pengecoh/distraktor yang ditolak karena tidak berfungsi sebelum dapat dimasukkan kedalam bank soal untuk tes-tes berikutnya.

KESIMPULAN

Pendidikan, baik di lingkungan sekolah maupun tidak, bertujuan untuk mengembangkan potensi dan kualitas berpikir seseorang. Pendidikan tinggi khususnya berupaya menghasilkan mahasiswa yang mampu berpikir kritis, inovatif, dan mengabdikan pada masyarakat. Mendengarkan merupakan salah satu keterampilan berbahasa yang harus diprioritaskan dalam kurikulum, karena sering digunakan dalam kehidupan sehari-hari. Evaluasi merupakan tugas penting pendidik yang meliputi penilaian dan evaluasi hasil dan proses pembelajaran. Tes pilihan ganda biasanya digunakan untuk mengevaluasi pemahaman siswa terhadap suatu mata pelajaran. Sebuah penelitian dilakukan untuk menilai kualitas ujian pilihan ganda pada mata kuliah menyimak pada mahasiswa PBSI Universitas Riau. Analisis butir soal dilakukan untuk menjamin mutu suatu tes dengan melihat tingkat kesukaran, daya pembeda, keefektifan alternatif jawaban, dan kemampuan membedakan nilai tinggi dan rendah. Penelitian ini menggunakan indikator reliabilitas tes, daya pembeda, tingkat kesukaran tes, dan peran pengecoh. Sampel diambil secara acak dari populasi sebanyak 119 mahasiswa. Data dianalisis menggunakan program komputer Anates versi 4.0.2.

Reliabilitas adalah ukuran apakah suatu instrumen, misalnya tes, dapat digunakan berulang kali untuk mengukur objek atau responden yang sama dengan hasil yang sama. Rumus KR digunakan untuk mengetahui reliabilitas instrumen tes pilihan ganda. Studi ini menemukan bahwa reliabilitas tes tersebut rendah, yang berarti tes tersebut tidak dapat digunakan berulang kali untuk mengukur objek atau responden yang sama dengan hasil yang sama. Tingkat daya pembeda pada tes didominasi dengan kategori sedang, hal ini menunjukkan bahwa kualitas daya pembeda pada tes pilihan ganda mampu membedakan siswa yang sudah menguasai materi dengan yang belum. Tingkat kesukaran soal ditemukan didominasi pada kategori sedang, hal ini menunjukkan bahwa soal tes tidak terlalu sulit dan tidak dapat dikatakan terlalu mudah. Tingkat keefektifan pengecoh pada tes ini dikategorikan diterima, direvisi, dan ditolak. Distraktor yang diterima efektif untuk digunakan pada pengujian selanjutnya, sedangkan distraktor yang ditolak harus diganti atau dibuang sebelum digunakan. Kesimpulannya, penelitian ini menemukan bahwa soal tes pilihan ganda untuk mata pelajaran mendengarkan perlu ditingkatkan, terutama dalam hal pertanyaan yang tidak relevan dan terlalu sulit, daya pembeda yang lemah, dan pengecoh yang tidak efektif. Perbaikan ini dapat membantu memastikan bahwa tes tersebut dapat diandalkan, valid, dan efektif dalam mengukur pemahaman siswa terhadap materi.

DAFTAR PUSTAKA

- Arif, M. (2014). Penerapan aplikasi anates bentuk soal pilihan ganda. *Jurnal Ilmiah Edutic: Pendidikan dan Informatika*, 1(1).
- Asrul, A, et al. 2022. *Evaluasi Pembelajaran*. Medan. Perdana Mulya Sarana Publishing.
- Fatimah, L. U., & Alfath, K. (2019). Analisis kesukaran soal, daya pembeda dan fungsi distraktor. *AL-MANAR: Jurnal Komunikasi dan Pendidikan Islam*, 8(2), 37-64.
- Halik, A. S., Mania, S., & Nur, F. (2019). Analisis butir soal Ujian Akhir Sekolah (UAS) mata pelajaran matematika pada tahun ajaran 2015/2016 SMP Negeri 36 Makassar. *Al Asma: Journal of Islamic Education*, 1(1), 11-17.
- Hermawan, Herry. 2012. *Menyimak Keterampilan Berkomunikasi yang Terabaikan*. Yogyakarta: Graha Ilmu.

- Khaerudin, K. (2016). Teknik Penskoran Tes Obyektif Model Pilihan Ganda. *Madaniyah*, 6(2), 183-200.
- Listiani, W., & Rachmawati, R. (2022). Transformasi Taksonomi Bloom dalam Evaluasi Pembelajaran Berbasis HOTS. *Jurnal Jendela Pendidikan*, 2(03), 397-402.
- Peraturan Pemerintah Nomor 4 Tahun 2022 tentang Standar Nasional Pendidikan
Permendikbudristek Nomor 8 Tahun 2024 tentang Standar Pendidikan Anak Usia Dini, Pendidikan Dasar, dan Pendidikan Menengah.
- Prihatin, Y. (2017). Problematika keterampilan menyimak dalam pembelajaran bahasa Indonesia. *Jurnal Sastronesia*, 5(3), 45-52.
- Rahmasari, D., & Ismiyati, I. (2016). Analisis Butir Soal Mata Pelajaran Pengantar Administrasi Perkantoran. *Economic Education Analysis Journal*, 5(1).
- Setiyawan, A. (2014). Faktor-Faktor yang Mempengaruhi Reliabilitas Tes. *AN NUR: Jurnal Studi Islam*, 6(2).
- Soulisa, I., Supratman, M., Rosfiani, O., Renaldi, R., Sopiah, S., Utomo, W. T & Sutisnawati, A. (2022). *Evaluasi Pembelajaran*.
- Sugiyono. 2017. *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D)* edisi ke-25. Bandung. ALFABETA.
- Suharsimi, Arikunto. 2013. *Dasar-Dasar Evaluasi Pendidikan Edisi 2*. Jakarta: PT Bumi Aksara.
- Undang-Undang Dasar Negara Republik Indonesia Tahun 1945
Undang-Undang Nomor 12 Tahun 2012 tentang Pendidikan Tinggi
- Yusup, F. (2018). Uji validitas dan reliabilitas instrumen penelitian kuantitatif. *Tarbiyah: Jurnal Ilmiah Kependidikan*, 7(1).