

Kualitas Butir Soal Ujian Akhir Semester Mata Kuliah Teori Akuntansi pada Program Studi Akuntansi

Nur Putra Rakhmadani¹, Dina Huriaty²

¹ Prodi Akuntansi STIE Pancasetia Banjarmasin, Kalimantan Selatan

² Prodi Pendidikan Matematika Universitas PGRI Kalimantan, Kalimantan Selatan

E-mail: putra.rakhmadani84@gmail.com¹, dina_huriaty@upk.ac.id²

Article History:

Received: 25 Januari 2025

Revised: 31 Januari 2025

Accepted: 07 Februari 2025

Keywords: Kualitas, Butir Soal, Akuntansi

Abstract: Penelitian ini bertujuan untuk mengetahui kualitas butir soal Ujian Akhir Semester Mata Kuliah Teori Akuntansi berdasarkan teori tes klasik, yaitu ditinjau dari tingkat kesukaran, daya pembeda, dan indeks reliabilitas tes. Metode dalam penelitian ini menggunakan penelitian eksploratif dengan pendekatan kuantitatif. Data yang digunakan adalah data sekunder, yaitu data respon ujian akhir semester mata kuliah teori akuntansi tahun akademik 2024-2025 mahasiswa program studi akuntansi STIE Pancasetia Banjarmasin, sebanyak 68 respon terhadap 30 butir soal bentuk pilihan ganda. Analisis dan estimasi terhadap indeks tingkat kesukaran, daya pembeda dan reliabilitas tes dilakukan menggunakan program Excell for Windows. Hasil penelitian menunjukkan ada tiga butir soal yang masuk dalam kategori sedang, tiga butir soal termasuk dalam kategori soal sukar, dan sisanya 24 butir soal mudah. Rata-rata indeks kesukaran butir secara keseluruhan termasuk pada kategori mudah. Berdasarkan indeks daya beda hanya dua butir soal, yaitu butir soal nomor 4 dan 12, yang termasuk pada kategori baik, sedangkan 28 butir yang lain pada kategori tidak baik. Secara keseluruhan rata-rata indeks daya beda menunjukkan termasuk pada kategori tidak baik. Hal ini berarti instrument ujian akhir semester tersebut tidak dapat membedakan kemampuan antara mahasiswa yang berkemampuan tinggi dengan mahasiswa yang berkemampuan rendah. Secara keseluruhan instrument ujian akhir semester memiliki indeks reliabilitas dibawah 0,7. Hasil penelitian ini direkomendasikan bagi pengajar untuk dapat melakukan analisis butir, terutama untuk soal berbentuk pilihan ganda. Hal ini perlu dilakukan untuk melihat seberapa tingkat kualitas butir soal yang telah dibuat. Butir soal yang telah memenuhi kriteria butir soal yang baik dapat menjadi bank soal yang dapat digunakan kembali untuk ujian berikutnya.

PENDAHULUAN

Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia Nomor 53 Tahun 2023 tentang Penjaminan Mutu Pendidikan Tinggi pada Paragraf ke 4 Pasal 26 ayat 1 dan 2 dinyatakan bahwa Standar penilaian merupakan kriteria minimal mengenai penilaian hasil belajar mahasiswa untuk mencapai standar kompetensi lulusan. Penilaian hasil belajar mahasiswa dilakukan secara valid, reliabel, transparan, akuntabel, berkeadilan, objektif, dan edukatif. Berdasarkan hal tersebut, instrumen evaluasi yang dibuat dosen seharusnya telah memenuhi kriteria tersebut. Evaluasi merupakan bagian dari rangkaian kegiatan dalam meningkatkan kualitas pendidikan. Suatu evaluasi semestinya didahului oleh kegiatan pengukuran (*measurement*) dan penilaian (*assessment*). Pengukuran hasil pembelajaran dapat dilakukan dengan menggunakan tes. Kualitas pendidikan ditentukan oleh kualitas pengukuran dan penilaian. Dengan demikian, kualitas pengukuran dan penilaian akan ditentukan juga oleh kualitas tes.

Tes sebagai alat ukur untuk memperoleh sejumlah informasi idealnya memiliki kualitas yang baik dan dikembangkan dan mengacu pada kurikulum yang digunakan. Hal ini berarti ketika mengembangkan tes, pengembang memperhatikan pada tujuan mata kuliah yang ingin dicapai seperti yang telah dituangkan dalam kurikulum yang berlaku. Selain itu setelah diadministrasikan, tes juga harus dianalisis untuk mengetahui karakteristik setiap butir soalnya.

Analisis butir soal bermanfaat untuk mengetahui suatu butir soal bermanfaat sesuai dengan yang diharapkan, mengevaluasi kemampuan mahasiswa dan dapat menjadi bahan diskusi di ruang kuliah, memberikan informasi kepada pengajar mengenai materi yang belum dikuasai, memberi masukan kepada pihak-pihak yang berwenang untuk mengevaluasi dan mengembangkan kurikulum, memperbaiki materi yang diujikan, serta meningkatkan kemampuan pengajar dalam menyusun soal (Brookhart & Nitko, 2011). Analisis butir soal dapat dilakukan untuk semua jenis tes dan bentuk soal.

Soal bentuk pilihan ganda merupakan bentuk soal yang seringkali digunakan pada ujian, mengingat berbagai kelebihan yang dimiliki oleh soal bentuk ini. Kelebihan soal bentuk pilihan ganda adalah bentuk soal yang dipilih pada pengujian skala besar dan hasilnya ingin segera diketahui. Selain itu, soal bentuk pilihan ganda memerlukan waktu yang relatif singkat dalam pelaksanaan, dapat diolah secara cepat, serta mudah dan objektif dalam penskorannya.

Namun seringkali pada pelaksanaannya sebuah tes tidak pernah dilakukan analisis terhadap kualitas butir soal yang akan diujikan. Mata kuliah Teori Akuntansi adalah mata kuliah wajib pada program studi Akuntansi di STIE Pancasetia Banjarmasin. Selama ini diketahui tidak pernah dilakukan analisis terhadap kualitas butir soal ujian akhir semester pada mata kuliah Teori Akuntansi, baik secara kualitatif maupun kuantitatif. Perlu diketahui bagaimana kualitas soal-soalnya, apakah memenuhi syarat validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh untuk soal yang baik. Soal dikatakan mempunyai kualitas yang baik apabila mempunyai validitas, reliabilitas, dan tingkat kesukaran yang sedang, memiliki daya pembeda serta pengecoh yang efektif dalam penggunaannya. Kevalidan alat evaluasi belum tentu dapat menjamin reliabilitas data selama penilaiannya masih menyimpan subjektivitas. Jadi, apabila syarat validitas dan reliabilitas terabaikan, rusaklah alat evaluasi tersebut.

Penelitian ini bertujuan untuk mengetahui kualitas butir soal Ujian Akhir Semester Mata Kuliah Teori Akuntansi berdasarkan teori tes klasik, yaitu ditinjau dari tingkat kesukaran, daya pembeda, dan indeks reliabilitas tes. Manfaat yang diperoleh dari hasil analisis butir soal adalah untuk mengetahui dan meningkatkan kualitas butir soal serta meningkatkan kompetensi pengampu mata kuliah dan hasil belajar mahasiswa.

LANDASAN TEORI

Analisis butir soal merupakan salah satu dari rangkaian evaluasi dalam suatu pembelajaran. Tujuannya adalah untuk mengkaji dan menelaah setiap butir soal agar diperoleh soal yang bermutu sebelum soal digunakan. Analisis butir soal juga dapat meningkatkan kualitas tes melalui revisi atau membuang soal yang tidak efektif, serta dapat digunakan untuk mengetahui informasi apakah *testee* sudah atau belum memahami materi yang telah diajarkan (Aiken, 1994: 63). Soal yang bermutu adalah soal yang dapat memberikan informasi setepat-tepatnya sesuai dengan tujuannya, diantaranya dapat menentukan mana yang sudah atau belum materi yang diajarkan telah dikuasai.

Analisis butir soal dapat dilakukan secara kualitatif dan kuantitatif. Analisis kualitatif mencakup analisis terhadap isi dan bentuknya, sedangkan analisis kuantitatif berhubungan dengan ciri-ciri statistiknya (Anastasi dan Urbina, 1997: 172). Analisis kualitatif terhadap isi dan konstruk dilakukan sebelum diujicobakan, sedangkan analisis kuantitatif mencakup pengukuran kesulitan butir soal dan diskriminasi butir soal, validitas soal dan reliabilitas instrument. Kualitas dari alat evaluasi diharapkan harus memperhatikan validitas dan reliabilitasnya. Analisis butir dilakukan untuk mendapatkan informasi tentang kualitas butir soal yang dibuat (Muliyani dan Huriaty, 2016).

Dua cara yang digunakan dalam penelaahan butir soal yaitu penelaahan soal secara kualitatif dan kuantitatif. Manfaat dilakukan analisis butir soal adalah: (1) dapat membantu para pengguna tes dalam evaluasi atas tes yang digunakan, (2) sangat relevan bagi penyusunan tes informal dan lokal seperti tes yang disiapkan, (3) mendukung penulisan butir soal yang efektif, (4) secara materi dapat memperbaiki tes di kelas, (5) meningkatkan validitas soal dan reliabilitas (Anastasi dan Urbina, 1997:172).

Linn dan Gronlund (1995: 3 16-318) menyatakan bahwa kegunaan analisis butir soal tidak hanya untuk meningkatkan kualitas butir soal, tetapi data analisis butir soal juga dapat digunakan sebagai dasar: (1) diskusi kelas efisien tentang hasil tes, (2) untuk kerja remedial, (3) untuk peningkatan secara umum pembelajaran di kelas, dan (3) untuk peningkatan keterampilan pada konstruksi tes.

Berbagai uraian di atas menunjukkan bahwa analisis butir soal bermanfaat untuk: (1) menentukan soal-soal yang cacat atau tidak berfungsi penggunaannya; dan (2) meningkatkan kualitas butir soal melalui tiga komponen analisis yaitu tingkat kesukaran, daya beda, dan efektivitas pengecoh soal.

Tingkat kesukaran didefinisikan sebagai proporsi peserta tes (*testee*) yang menjawab suatu butir soal dengan benar. Tingkat kesukaran dilambangkan dengan p . Besarnya indeks kesukaran berkisar antara 0,00 sampai dengan 1,00. Jika suatu butir soal mempunyai $p = 0,00$, berarti butir soal tersebut sangat sukar, dan jika suatu butir soal mempunyai harga $p = 1,00$, berarti item sangat mudah, karena setiap peserta tes dapat menjawab dengan benar. Hal ini dapat disimpulkan bahwa semakin tinggi harga p suatu butir maka butir soal tersebut semakin mudah, dan sebaliknya.

Daya beda adalah kemampuan suatu butir soal dalam membedakan antara *testee* yang mempunyai kemampuan tinggi dan *testee* yang mempunyai kemampuan rendah. Indeks daya beda berkisar antara -1 sampai dengan +1. Suatu butir soal dikatakan baik jika daya beda mempunyai harga minimal 0,30 (Kartowagiran, 2009).

Reliabilitas adalah derajat konsistensi antara dua skor hasil pengukuran pada objek yang sama yang menggunakan alat ukur dan skala yang berbeda. Tes dikatakan reliabel jika skor amatan memiliki korelasi yang tinggi dengan *true-score* nya (Allen & Yen, 1979). Dapat disimpulkan bahwa tes yang reliabel adalah tes yang memiliki kestabilan hasil ukur meskipun dilakukan pengukuran berulang.

Indeks reliabilitas suatu tes merupakan nilai estimasi yang diperoleh dari perhitungan

korelasi, dimana indeks korelasi berkisar dari -1 hingga +1. Indeks korelasi mendekati +1 menunjukkan reliabilitas tinggi dan sebaliknya. Indeks reliabilitas tes yang dapat diterima adalah minimal 0,70 (Allen & Yen, 1979), tetapi untuk indeks 0,65 pada soal bentuk uraian masih dapat diterima. Sementara itu menurut Ghazali (2016), suatu instrumen penelitian dikatakan dapat diandalkan (reliable) apabila nilai *Cronbach's Alpha* $\geq 0,60$.

Mata kuliah teori akuntansi di program studi Akuntansi bertujuan untuk membekali dan memberikan pemahaman kepada mahasiswa tentang perkembangan dan pembentukan teori akuntansi. Mata kuliah ini berisi tentang ruang lingkup akuntansi, akuntansi dan *double entry*, perkembangan ilmu akuntansi, sejarah akuntansi di Indonesia, elemen struktur teori akuntansi, perumusan teori akuntansi, konsep dasar laporan keuangan, standar akuntansi keuangan, dan kasus-kasus yang berkaitan dengan materi.

Disajikan dalam 3 SKS, setelah mempelajari teori akuntansi, mahasiswa diharapkan mampu menguasai kerangka teoritis akuntansi, memahami perkembangan teori dalam akuntansi dan kerangka konseptual akuntansi, dan mampu memahami hubungan antara teori akuntansi dengan praktek dan kebijakan akuntansi.

METODE PENELITIAN

Metode dalam penelitian ini menggunakan penelitian eksploratif dengan pendekatan kuantitatif. Data yang digunakan adalah data sekunder, yaitu data respon ujian akhir semester mata kuliah teori akuntansi tahun akademik 2024-2025 mahasiswa program studi akuntansi STIE Pancasetia Banjarmasin, sebanyak 68 respon terhadap 30 butir soal bentuk pilihan ganda. Analisis terhadap indeks tingkat kesukaran, daya pembeda dan estimasi reliabilitas tes dilakukan menggunakan program *Excell for Windows*.

Saifuddin Azwar (2003: 134) menyatakan bahwa indeks kesukaran butir adalah rasio penjawab butir dengan benar dan banyaknya penjawab butir. Proporsi menjawab benar p (*proportion correct*) adalah indeks kesukaran soal yang paling sederhana dan sering digunakan dalam menentukan besaran indeks. Rumus untuk menentukan besarnya indeks kesukaran butir secara matematis dirumuskan sebagai berikut.

$$p = \frac{B}{N}$$

Keterangan:

p : indeks kesukaran butir

B : banyaknya *testee* yang menjawab benar suatu butir

N : banyaknya *testee*

Di bawah ini adalah tabel kategori suatu butir soal berdasarkan harga p .

Tabel 1. Kategori Butir Soal Berdasarkan IndeksTingkat Kesukaran

Tingkat kesukaran (p)	Kategori
$0,0 < p < 0,3$	Sukar
$0,3 \leq p \leq 0,7$	Sedang

$0,7 < p < 1,0$	Mudah
-----------------	-------

Daya beda adalah perbedaan antara proporsi kelompok atas yang menjawab benar suatu butir dengan proporsi kelompok bawah yang menjawab benar butir soal tersebut. Pengertian ini diformulasikan oleh Brennan (1972) sebagaimana dikutip Yen W.M dalam Encyclopedia of Educational Research dalam Lababa (2018), dapat dituliskan sebagai berikut.

$$DB = p_A - p_B$$

Formulasi ini dapat digunakan untuk menghitung daya beda butir soal dalam bentuk pilihan ganda. Analisis terhadap daya pembeda butir soal mengacu pada Tabel 2 berikut ini.

Tabel 2. Kategori Butir Soal Berdasarkan Indeks Daya Beda

Daya beda	Kategori
$DB < 0,30$	Tidak Baik
$DB \geq 0,30$	Baik

Reliabilitas suatu instrumen dapat dilihat dengan menggunakan pendekatan *internal consistency*. Pada penelitian ini menggunakan formulasi *Alpha Cronbach*, sebagai berikut.

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum \sigma_x^2}{\sigma_t^2} \right)$$

Keterangan:

r_{11} : reliabilitas

r_{11} : banyaknya butir soal

$\sum \sigma_x^2$: Jumlah varians skor butir

σ_t^2 : Varians total

(Retnawati, Heri: 2015).

HASIL DAN PEMBAHASAN

Hasil analisis terhadap 30 butir soal didapatkan indeks tingkat kesukaran, daya pembeda butir soal dan reliabilitas tes. Hasil ini dapat dilihat pada tabel berikut ini.

Tabel 3. Hasil Analisis Butir Soal

Nomor Butir Soal	Tingkat Kesukaran		Daya Beda	
	Indeks	Kategori	Indeks	Kategori
1	0,99	Mudah	0,03	Tidak Baik
2	1,00	Mudah	0,00	Tidak Baik
3	1,00	Mudah	0,00	Tidak Baik
4	0,32	Sedang	0,35	Baik

5	0,97	Mudah	0,06	Tidak Baik
6	0,79	Mudah	0,24	Tidak Baik
7	0,25	Sukar	0,21	Tidak Baik
8	0,94	Mudah	0,00	Tidak Baik
9	0,99	Mudah	0,03	Tidak Baik
10	0,78	Mudah	0,26	Tidak Baik
11	0,81	Mudah	0,21	Tidak Baik
12	0,38	Sedang	0,35	Baik
13	0,24	Sukar	0,24	Tidak Baik
14	0,87	Mudah	0,21	Tidak Baik
15	0,97	Mudah	0,06	Tidak Baik
16	0,96	Mudah	0,03	Tidak Baik
17	0,91	Mudah	0,12	Tidak Baik
18	0,19	Sukar	0,26	Tidak Baik
19	0,99	Mudah	0,03	Tidak Baik
20	0,85	Mudah	0,12	Tidak Baik
21	1,00	Mudah	0,00	Tidak Baik
22	0,93	Mudah	0,15	Tidak Baik
23	0,99	Mudah	0,03	Tidak Baik
24	0,71	Mudah	0,24	Tidak Baik
25	0,68	Sedang	0,12	Tidak Baik
26	0,88	Mudah	0,12	Tidak Baik
27	0,97	Mudah	0,06	Tidak Baik
28	0,96	Mudah	0,09	Tidak Baik
29	0,96	Mudah	0,03	Tidak Baik
30	0,97	Mudah	0,06	Tidak Baik
Rata-Rata	0,81	Mudah	0,12	Tidak Baik
Indeks Reliabilitas instrumen			0,65	

Berdasarkan Tabel 3, hasil analisis terhadap 30 butir soal menunjukkan indeks reliabilitas instrumen adalah 0,65. Mengacu pada Allen & Yen (1979), instrumen UAS Mata Kuliah Teori Akuntansi ini masih di bawah 0,70, namun menurut Ghazali (2016), indeks reliabilitas 0,65 masih dapat diterima. Hasil ini menunjukkan sejauh mana konsistensi hasil pengukuran yang jika dilakukan berulang-ulang akan memperoleh hasil yang relatif sama dengan instrumen UAS yang sama.

Indeks tingkat kesukaran butir berdasarkan kategori, dari hasil analisis pada Tabel 3 diperoleh dari 30 butir soal terdapat 24 butir (80%) yang termasuk dalam kategori mudah, terdapat 3 butir (10%) dengan kategori sedang, yaitu butir soal nomor 4, 12, dan 25; dan 3 butir dengan kategori sukar atau 10%, yaitu butir soal nomor 7, 13, dan 18. Rata-rata indeks kesukaran butir soal UAS sebesar 0,81 termasuk dalam kategori mudah, artinya butir-butir soal UAS Mata Kuliah Teori Akuntansi yang diadministrasikan tidak bisa memberikan informasi tentang kemampuan mahasiswa, karena tingkat kesukaran butir-butir soal berada di bawah kemampuan mahasiswa sebenarnya.

Analisis terhadap indeks daya beda 30 butir soal menunjukkan hanya 2 butir (6,7%) dalam kategori baik yaitu butir soal nomor 4 dan 12. Hasil ini menunjukkan bahwa butir 4 dan 12 dapat

membedakan mahasiswa yang kemampuannya tinggi dengan mahasiswa yang kemampuannya rendah. Rata-rata indeks daya pembeda butir secara keseluruhan sebesar 0,12. Indeks ini berada di bawah 0,3, artinya butir-butir soal pada instrumen UAS Mata Kuliah Teori Akuntansi yang diadministrasikan tidak dapat membedakan mahasiswa yang kemampuannya tinggi dengan mahasiswa yang memiliki kemampuan rendah.

Secara keseluruhan hanya ada dua butir soal yang memenuhi kriteria soal yang baik, yaitu no 4 dan 12, yang dapat dilihat pada gambar di bawah ini.

4. Berikut ini adalah proses akuntansi, kecuali
- Mark only one oval.*
- ☐ pencatatan
- ☐ penggolongan
- ☐ perencanaan
- ☐ pelaporan

Gambar 1. Butir Soal

Pada butir soal nomor 4, merupakan pertanyaan tentang yang bukan proses akuntansi. Jawaban yang benar adalah “perencanaan”. Butir soal nomor 4 ini dijawab benar oleh 22 peserta UAS.

12. Pembelian perlengkapan secara tunai dalam persamaan akuntansi akan :
- Mark only one oval.*
- ☐ Mengurangi aset dan menambah modal
- ☐ Mengurangi utang dan mengurangi aset
- ☐ Menambah utang dan menambah modal
- ☐ Menambah aset dan mengurangi aset

Gambar 2. Butir Soal

Sedangkan butir soal nomor 15 merupakan soal tentang persamaan akuntansi. Jawaban yang tepat adalah “menambah aset dan mengurangi aset”. Pada butir soal ini dijawab benar oleh 26 peserta UAS.

Melihat nilai rata-rata hasil ujian akhir semester pada mata kuliah Teori Akuntansi, mahasiswa dapat menjawab dengan benar 24 sampai dengan 25 butir soal, dengan rata-rata nilai 80,74 atau setara nilai A (4). Namun, berdasarkan hasil analisis terhadap butir soal ternyata tidak dapat memberikan informasi tentang tingkat kemampuan mahasiswa sesungguhnya, disebabkan instrumen UAS Mata Kuliah Teori Akuntansi yang diadministrasikan termasuk dalam kategori mudah. Butir-butir soal yang ditolak bukan berarti tidak dapat digunakan. Butir-butir soal tersebut dapat direview kembali dan diperbaiki untuk selanjutnya diujicobakan pada kelompok *testee* yang lebih heterogen. Hal ini dapat dilakukan mengingat analisis dengan teori tes klasik sangat tergantung dengan kemampuan sampel uji coba yang dipilih (*dependent sampling*) (Muliyani dan Huriati, 2016).

KESIMPULAN

Berdasarkan hasil analisis dan pembahasan dapat disimpulkan bahwa kualitas butir soal ujian akhir semester pada mata kuliah teori akuntansi ditinjau dari indeks tingkat kesukaran, menunjukkan ada tiga butir soal yang masuk dalam kategori sedang, tiga butir soal termasuk dalam kategori soal sukar, dan sisanya 24 butir soal mudah. Rata-rata indeks kesukaran butir secara

keseluruhan termasuk pada kategori mudah. Berdasarkan indeks daya beda hanya dua butir soal, yaitu butir soal nomor 4 dan 12 yang termasuk pada kategori baik, sedangkan 28 butir yang lain pada kategori tidak baik. Secara keseluruhan rata-rata indeks daya beda menunjukkan termasuk pada kategori tidak baik. Hal ini berarti instrument ujian akhir semester tersebut tidak dapat membedakan kemampuan antara mahasiswa yang berkemampuan tinggi dengan mahasiswa yang berkemampuan rendah. Secara keseluruhan instrument ujian akhir semester memiliki indeks reliabilitas dibawah 0,7. Hasil penelitian ini direkomendasikan bagi pengajar untuk dapat melakukan analisis butir, terutama jika soal berbentuk pilihan ganda. Hal ini perlu dilakukan untuk melihat seberapa kualitas butir soal yang telah dibuat pengajar. Butir soal yang telah memenuhi kriteria butir soal yang baik dapat menjadi bank soal yang dapat digunakan kembali untuk ujian berikutnya.

DAFTAR REFERENSI

- Aiken, L. R. (1987). *Assessment of Intellectual Functioning*. Massachussetts: Allyn and Bacon Inc. -----, (1994). *Psychological Testing and Assessment*. Boston: Allyn and Bacon, Inc.
- Allen, M. J., & Yen, W. M. (1979). *Introduction to Measurement Theory*. Monterey, California: Brookd/Cole Publishing Company.
- Anastasi, Anne dan Urbina, Susana. (1997). *Psychological Testing*. (Seventh Edition). New Jersey: Prentice-Hall, Inc.
- Brookhart, S. M., & Nitko, A. J. (2011). *Educational Assessment of Students*. Pearson Higher Ed.
- Ghozali, I. (2016). *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 23*. Edisi 8. Semarang: Badan Penerbit Universitas Diponegoro
- Kartowagiran, B. (2009). *Pengantar Teori Tes Klasik (TTK)*.
- Lababa, J. (2018). Analisis butir soal dengan teori tes klasik: Sebuah pengantar. *Jurnal Ilmiah Iqra'*, 2(2).
- Linn, R. L. dan Gronlund. (1995). *Measurement and Assessment in Teaching*. Columbus, Ohio: Merrill Prentice Hall Education
- Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia Nomor 53. (2023). *Penjaminan Mutu Pendidikan Tinggi*.
- Retnawati, Heri. (2015). *Analisis Kuantitatif Instrumen Penelitian (Panduan Peneliti, Mahasiswa, dan Psikometrian)*. Yogyakarta: Parama Publishing.
- Saifuddin Azwar. (2003). *Tes Prestasi: Fungsi dan Pengembangan Pengukuran Prestasi Belajar*. Yogyakarta: Pustaka Pelajar.
- T. Muliyani dan D. Huriaty. (2016). Pengembangan Instrumen Tes Geometri dan Pengukuran pada Jenjang SMP. *Jurnal Math Didactic*. Vol. 2, No. 2, pp. 91–98.