

ANALISIS KUALITAS SOAL UJIAN SEKOLAH BERBASIS NASIONAL MATA PELAJARAN IPA SMP NEGERI DI KABUPATEN LUWU TIMUR BERDASARKAN TEORI RESPON BUTIR

Ishak

Kaharuddin Arafah

Ruslan

Ishaknewarchi@gmail.com

Abstrak: Penelitian ini bertujuan untuk mendeskripsikan kualitas Soal Ujian Sekolah Berbasis Nasional Mata Pelajaran IPA SMP Negeri di Luwu Timur. Penekanan penelitian ini pada aspek Validitas Isi, Model Logistik 1 Parameter, 2 Parameter, dan 3 Parameter. Penelitian ini merupakan penelitian kuantitatif deskriptif. Subjek penelitian adalah seluruh respon peserta didik terhadap soal USBN Mata Pelajaran IPA SMP Negeri di Luwu Timur sebanyak 4.094 lembar jawaban. Pengumpulan data dilakukan dengan dokumentasi. Teknik analisis data yang digunakan adalah kuantitatif deskriptif. Hasil penelitian menunjukkan bahwa soal USBN Mata Pelajaran IPA SMP Negeri di Luwu Timur belum memenuhi validitas isi. Untuk Model Logistik 1 Parameter, yang meninjau aspek tingkat kesukaran terdapat 80% butir dengan kualitas baik, dan 20% butir dengan kualitas tidak baik. Untuk Model Logistik 2 Parameter, yang meninjau aspek tingkat kesukaran dan daya beda terdapat 42, 86% butir dengan kualitas baik, dan terdapat 57,14% butir dengan kualitas tidak baik. Untuk Model Logistik 3 Parameter, yang meninjau aspek tingkat kesukaran, daya beda, dan tingkat kesukaran terdapat 28,57% butir dengan kualitas baik, dan terdapat 71,43% butir dengan kualitas tidak baik.

Kata Kunci: Teori Respon Butir

***Abstract:** This study aims to describe the quality of the National Junior High School Examination Science School Subject Questions in East Luwu in terms of Content Validity, Logistics Model 1 Parameter, 2 Parameters, and 3 Parameters. This research is descriptive quantitative research. The research subjects were all students' responses to the questions of the State Junior High School Science in East Luwu as many as 4,094 answer sheets. Data collection is done by documentation. The data analysis technique used is quantitative descriptive. The results of the study show that the problem of the USB Science in the State Middle School in East Luwu has not fulfilled the content validity. For Logistics Models 1 Parameters, which review the difficulty level, there are 80% of items of good quality, and 20% of items of poor quality. For the Logistics Model 2 parameters, which reviewed aspects of the level of difficulty and power differences, there were 42, 86% of the items with good quality, and there were 57.14% of items with poor quality. For the Logistics Model 3 parameters, which review aspects of the level of difficulty, power difference, and the level of difficulty there are 28.57% of items of good quality, and there are 71.43% of items with poor quality.*

KeyWords: Item Response Theory

1. Pendahuluan

Untuk mengukur dan menilai pelaksanaan USBN diperlukan sebuah instrumen sebagai perangkat ukur kemampuan kognitif peserta didik yaitu berupa tes hasil belajar. Salah satu bentuk tes yang paling sering digunakan adalah tes pilihan ganda. Tes ini dianggap sangat sesuai digunakan untuk mengukur cakupan materi yang luas dan sangat efektif untuk peserta tes yang sangat banyak, yang mengharuskan sesegera mungkin untuk diketahui hasilnya, seperti pada kegiatan USBN SMP yang diselenggarakan di Kabupaten Luwu Timur..

Dalam prosesnya, fakta dilapangan bahwa perakitan soal USBN Mata Pelajaran IPA untuk tingkat SMP di Kabupaten Luwu Timur dirakit oleh guru-guru Mata Pelajaran IPA yang penunjukannya oleh Dinas Pendidikan dengan mengacu pada kisi-kisi yang telah ditetapkan oleh BNSP sesuai dengan kriteria pencapaian standar kompetensi lulusan (SKL), standar isi, dan kurikulum yang berlaku.

Untuk mendapatkan alat ukur yang berkualitas maka perlu dilakukan analisis butir soal secara teoretis dan analisis secara empiris. Kedua jenis analisis soal ini, para perakit soal hanyalah menganalisis soal secara teoretis dan mengabaikan analisis secara empiris. Sehingga kualitas soal USBN SMP Negeri di Kabupaten Luwu Timur jika ditinjau berdasarkan analisis secara empiris belum bisa dipastikan kelayakannya.

Hal tersebut sejalan dengan data yang didapatkan dari hasil wawancara dengan salah seorang guru perakit soal mengatakan “ Perakitan soal USBN hanya berpatokan pada kisi-kisi yang dikeluarkan oleh BNSP dengan melihat dimensi indikatornya saja, adapun untuk mengetahui tingkat kesukaran, daya beda, dan jawaban pengecoh tidak menjadi pertimbangan karena soal tersebut diambil dari berbagai sumber yang ada”. Dari pernyataan ini dapat diartikan bahwa mekanisme perakitan soal USBN belum mengikuti prosedur baku pembuatan tes yang standar.

Belum adanya penelitian terkait kualitas soal yang digunakan selama ini, sementara USBN adalah salah satu variabel penting bagi penentuan kelulusan peserta didik untuk tingkat satuan pendidikan. Sehingga penelitian ini difokuskan pada analisis kualitas butir soal USBN Mata Pelajaran IPA SMP.

Dari berbagai permasalahan yang dikemukakan tersebut, maka sangat penting untuk melakukan kajian yang lebih mendalam sebagai upaya mengungkap kualitas Soal USBN SMP Negeri di Kabupaten Luwu Timur yang digunakan selama ini. Salah satu solusi yang tepat untuk dilakukan adalah menganalisis kualitas butir soal USBN dengan menerapkan Teori Modern atau Teori Respon Butir. Pemilihan teori respon butir didasarkan pada klasifikasi dan tujuan tes serta kelebihan yang dimiliki oleh teori tersebut, sekaligus sebagai informasi baru bagi rekan guru yang ada di luwu timur dalam menganalisis soal dengan konsep teori respon butir.

Berdasarkan latar belakang di atas, maka akan dikemukakan beberapa masalah dalam penelitian ini yaitu:

1. Bagaimanakah kualitas soal USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur ditinjau dari Validitas Isi?
2. Bagaimanakah kualitas soal USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur ditinjau dari Model Logistik 1 Parameter?
3. Bagaimanakah kualitas soal USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur ditinjau dari Model Logistik 2 Parameter?
4. Bagaimanakah kualitas soal USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur ditinjau dari Model Logistik 3 Parameter?

Dari uraian permasalahan di atas, tujuan yang diharapkan dapat dicapai dalam penelitian ini yaitu untuk mengetahui:

1. Kualitas soal USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur ditinjau dari Validitas Isi.
2. Kualitas soal USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur ditinjau dari Model Logistik 1 Parameter.
3. Kualitas soal USBN mata pelajaran IPA SMP Negeri di Kabupaten Luwu Timur ditinjau dari Model Logistik 2 Parameter.
4. Kualitas soal USBN mata pelajaran IPA SMP Negeri di Kabupaten Luwu Timur ditinjau dari Model Logistik 3 Parameter.

Manfaat yang diharapkan dari pelaksanaan penelitian ini adalah sebagai berikut:

1. Bagi guru, sebagai bahan referensi dalam merakit soal yang berkualitas dan sesuai dengan prosedur baku dalam pembuatan soal yang standar dan memenuhi prinsip-prinsip pengukuran serta format penulisan soal.
2. Bagi sekolah, tersedianya bank soal yang sudah diketahui kualitasnya yang tentunya dapat dipergunakan pada kegiatan penilaian hasil belajar lainnya.
3. Bagi MGMP IPA SMP di Kabupaten Luwu Timur, sebagai bahan masukan untuk perakitan soal USBN Mata Pelajaran IPA SMP di Kabupaten Luwu Timur untuk tahun-tahun berikutnya.
4. Bagi Dinas Pendidikan Kabupaten Luwu Timur, sebagai dasar dalam menentukan kebijakan untuk perakitan soal USBN Mata Pelajaran IPA SMP di tahun berikutnya (Tim Perakit Soal).
5. Bagi peneliti, sarana untuk mengaplikasikan secara langsung bekal ilmu yang telah didapatkan sebagai mahasiswa yang diberi amanah dan dibiayai oleh Pemerintah Daerah Kabupaten Luwu Timur dengan menggunakan pendekatan teori respon butir dalam menganalisis kualitas soal USBN Mata Pelajaran IPA SMP tahun pelajaran 2017/2018 di Kabupaten Luwu Timur.

2. Metode Penelitian

Jenis penelitian ini adalah penelitian deskriptif dengan pendekatan kuantitatif. Subjek penelitian ini adalah semua data hasil jawaban peserta didik USBN Mata Pelajaran IPA SMP Negeri Di Kabupaten Luwu Timur Tahun Pelajaran 2017/2018 sebanyak 4.094 peserta didik. Namun setelah dilakukan pengumpulan data berupa dokumen lembar jawaban peserta didik di SMP Negeri yang tersebar di Kabupaten Luwu Timur banyaknya dokumen yang diperoleh hanya 2.282 lembar jawaban.

Teknik analisis data dalam penelitian ini adalah analisis data secara teoritik dan empirik. Untuk analisis data secara teoritik dilakukan dengan penelaan butir soal dengan meminta pertimbangan ahli (*Expert Judgement*). Sedangkan analisis secara empirik dilakukan dengan menelaah karakteristik setiap butir soal atau item berdasarkan teori respon butir. Analisis ini dilakukan dengan menggunakan bantuan aplikasi program (*soft ware*) komputer yaitu Program R i386 Vr 3.5.1.

1.1 Validitas Isi

Validitas isi mencakup hal-hal yang berkaitan dengan soal dari aspek materi, konstruksi dan bahasa dengan berdasarkan telaah oleh dua orang ahli (*expert judgement*).

Dimana kesepakatan antar dua pakar memberikan suatu metode untuk menentukan validitas isi secara menyeluruh (*Overal*) berdasarkan judgements of expert, yaitu mengevaluasi butir soal dengan

menggunakan skala empat (Model Gregory)

Instrumen yang mempunyai validasi isi $> 0,75$ dapat dinyatakan hasil pengukuran atau intervensi yang dilakukan oleh kedua pakar adalah valid (Ruslan, 2009). Penentuan jumlah butir soal untuk sel A, sel B, sel C, dan sel D berdasarkan model Gregory, dengan nilai koefesien validasi isi sebagai berikut:

$$\text{Validitas Isi} = \frac{D}{A + B + C + D}$$

1.2 Analisis Berdasarkan Teori Respon Butir

a. Independen Lokal

Independen lokal menginginkan butir-butir soal independen, artinya tidak berkorelasi tinggi antara butir satu dengan butir yang lainnya. Untuk mengestimasi hal tersebut dilakukan pembagian dua kelompok butir ganjil dan genap, kemudian parameter butir tersebut dikorelasikan. Jika $r_{ab} < 0.8$ maka asumsi independen lokal terpenuhi.

b. Unidimensi

Untuk mengetahui bahwa tes hanya mengandung satu komponen dominan yang mengukur kemampuan peserta. Uji unidimensi dilakukan dengan estimasi model logistik 1PL, 2PL dan 3PL dengan *software R* Program. Hasil estimasi uji unidimensi membandingkan nilai *p-value* dengan $\alpha = 0,05$. Kriteria uji unidimensi terpenuhi jika *p-value* observasi lebih kecil dari $\alpha = 0,05$.

c. Uji kecocokan Model

Pemilihan model parameter yang cocok dengan melihat nilai terendah dari BIC (*Bayesian Information Criteria*) pada masing-masing model.

Menurut Whittaker, Chang, dan Dodd (dalam Ali Ridho, 2014:7) sampai dengan saat ini ada beberapa kriteria kecocokan data yang dapat digunakan dalam menentukan model mana yang paling menjelaskan data, yaitu: *Likelihood Ratio Test* (LRT) atau perbedaan $-2\log$ likelihood, *Akaike's Information Criterion* (AIC), *Bayesian Information Criterion* (BIC), *Cross-validation Log Likelihood* (CVLL), dan *Deviance Information Criterion* (DIC). Dalam hal ini dinyatakan bahwa BIC efektif untuk memilih indeks kecocokan data untuk sampel yang besar ($n > 1000$).

d. Estimasi 1PL

Parameter tingkat kesukaran (b_i) bervariasi dari -2 sampai +2. Nilai mendekati -2 menunjukkan bahwa butir soal tersebut sangat mudah. Nilai mendekati +2 menunjukkan bahwa butir soal tersebut sangat sulit (Hambleton dan Swaminathan, 1985).

e. Estimasi 2 PL :

Parameter daya beda (a_i) didefinisikan pada skala $(-\infty, +\infty)$. Semakin kecil koefisien daya beda menunjukkan soal dengan daya beda jelek, semakin besar koefisien daya beda menunjukkan soal yang mempunyai daya beda yang baik. (Hambleton dan Swaminathan, 1985).

f. Estimasi 3 PL

Untuk jawaban tebakan probabilitas peluang menebak, bergantung kepada banyaknya pilihan. Soal yang ideal adalah soal yang memiliki faktor quessing mendekati nol di bawah 0,25 ($c_i < 0,25$), (Hullin, 1983).

3. Hasil Penelitian dan Pembahasan

3.1 Deskripsi Kualitas Tes USBN Mata Pelajaran IPA SMP Negeri Tahun Pelajaran 2017/2018 berdasarkan Validitas Isi

Pertanyaan penelitian pertama yang ditinjau untuk mengidentifikasi kualitas tes USBN yaitu validitas isi. Validitas isi tes USBN yang dimaksud merupakan gambaran sejauh mana kesesuaian antara butir-butir tes USBN dengan standar kompetensi, kompetensi dasar, indikator dan materi yang tercantum dalam SKL pada Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur Tahun Pelajaran 2017/2018

Hasil validasi isi dari dua pakar terhadap 35 butir tes USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur Tahun Pelajaran 2017/2018 tersebut di atas diperoleh informasi bahwa koefisien validasi isi tes sebesar 0,65. Oleh karena koefisien validasi lebih kecil dari 0,75 maka dapat diputuskan bahwa tes tersebut tidak memenuhi validitas isi berdasarkan telaah kedua pakar, sehingga ada beberapa butir yang perlu untuk ditindak lanjuti (Revisi/perbaiki) berdasarkan telaah oleh dua pakar

3.2 Deskripsi Soal USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur TP. 2017/2018 berdasarkan Model Logistik 1 Parameter.

Pertanyaan penelitian kedua terkait dengan kualitas Tes USBN Mata Pelajaran IPA ditinjau berdasarkan parameter tingkat

kesukaran butir tes (b_i) dengan nilai daya pembeda semua butir adalah 0,45 dan tebakan sama dengan nol. Sesuai pendapat Baker (2001: 22), menyatakan bahwa parameter tingkat kesukaran dilambangkan dengan b , didefinisikan sebagai titik pada kemampuan skala dimana kemungkinan respons yang benar terhadap item adalah 0,5.

Setyawarno (2016), membedakan tingkat kesukaran tes ke dalam lima kategori dengan melihat nilai Δ atau *threshold* yaitu butir yang memiliki nilai koefisien ($b > 2$) termasuk kategori sangat sukar, ($1 < b \leq 2$) termasuk kategori sukar, ($-1 \leq b \leq 1$) disebut termasuk kategori sedang, ($-1 > b \geq -2$) disebut sebagai kategori mudah, ($b < -2$) disebut sebagai kategori sangat mudah.

Berdasarkan pengkategorian tersebut diperoleh data empirik yang menunjukkan bahwa terdapat 28 atau 80% butir dengan kualitas baik dan terdapat 7 atau 20% butir dengan kualitas tidak baik. Sesuai pendapat Hambleton, Swaminathan, dan Rogers (1991: 13) mengemukakan bahwa nilai b_i atau indeks tingkat kesukaran butir yang baik berkisar dari (-2) – (+2). Nilai yang mendekati garis negatif menunjukkan bahwa butir tersebut terlalu mudah, begitupula semakin mendekati garis positif menunjukkan butir soal tersebut semakin sulit.

Butir tes yang tidak baik ini disarankan untuk dilakukan perbaikan atau revisi dengan memperhatikan tingkat kesukarannya, sehingga diperoleh suatu tes yang berkualitas.

Berdasarkan fungsi tes tersebut jika dikaitkan dengan tingkat

kesukaran maka butir tes yang digunakan harus proporsional. Sejalan dengan pendapat Zainal Arifin (2017: 70), proporsional artinya suatu instrumen harus memiliki tingkat kesukaran antara sulit, sedang, dan mudah yang ideal.

3.3 Kualitas soal USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur ditinjau dari Model Logistik 2 Parameter

Hasil analisis butir soal dengan Model Logistik 2 Parameter dengan melihat peluang peserta tes untuk menjawab dengan benar suatu butir soal ditentukan oleh dua karakteristik butir, yaitu tingkat kesulitan butir soal dan daya beda soal dan untuk faktor tebakan masih diasumsikan sama dengan nol atau tidak ada tebakan.

Temuan empirik menunjukkan bahwa kualitas butir tes USBN untuk model logistik 2 parameter terdapat 15 atau 42, 86% butir dengan kualitas baik, dan terdapat 20 atau 57,14% butir dengan kualitas tidak baik. Pengkategorian butir tes yang tidak baik ini diperoleh dengan melihat kesesuaian nilai parameter daya beda dan tingkat kesukaran yang selanjutnya dilakukan perbaikan/revisi terhadap butir tes tersebut

Adapun parameter a_i (daya beda) didefinisikan pada skala $(-\infty, +\infty)$. Nilai yang negatif menunjukkan bahwa butir yang bersangkutan harus dibuang. Nilai yang positif menunjukkan bahwa tes tersebut berfungsi dengan baik dalam mengukur kemampuan yang hendak diukur (Hambleton & Swaminathan,

1986: 37). Semakin kecil koefisien daya beda menunjukkan soal dengan daya beda jelek, semakin besar koefisien daya beda menunjukkan soal yang mempunyai daya beda yang baik. Namun daya beda yang terlalu rendah atau terlalu tinggi tidak dapat dijadikan tolak ukur dalam membedakan kemampuan peserta tes secara signifikan. Nilai daya beda yang baik berkisar antara 0 sampai dengan 2.

Butir tes yang tidak baik ini disarankan untuk dilakukan perbaikan atau revisi dengan memperhatikan tingkat kesukarannya, sehingga diperoleh suatu tes yang berkualitas.

3.4 Kualitas soal USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur ditinjau dari Model Logistik 3 Parameter

Pada Model 3 PL peluang peserta tes untuk menjawab dengan benar suatu butir tes ditentukan oleh tiga karakteristik butir, yaitu tingkat kesukaran butir tes, daya beda, dan peluang menebak. Model 3 parameter mempertimbangkan faktor tebakan (c_i), dengan alasan bahwa peserta tes akan dapat menjawab butir dengan benar melalui faktor tebakan namun kemungkinan respon menjawab dengan benar bernilai rendah.

Berdasarkan ketiga karakteristik tersebut kualitas butir tes USBN untuk model logistik 3 parameter temuan empirik menunjukkan bahwa terdapat 10 atau 28, 57% butir dengan kualitas baik dan terdapat 25 atau 71,43% butir dengan kualitas tidak baik. Pengkategorian butir tes yang tidak baik ini diperoleh dengan melihat kesesuaian nilai parameter daya beda,

tingkat kesukaran, dan jawaban tebakan yang selanjutnya dilakukan perbaikan/revisi terhadap butir tes tersebut.

Menurut Hullin (Heri Retnawati, 2014: 18) untuk jawaban tebakan pada suatu butir tes, nilai c_i ini berkisar antara 0 dan 1. Suatu butir dikatakan baik jika nilai c_i tidak lebih dari $\frac{1}{k}$ dengan k banyaknya opsi atau pilihan. Karena banyaknya pilihan pada Tes USBN Mata Pelajaran IPA adalah 4 pilihan jawaban untuk setiap butir tesnya maka batasan maksimum faktor tebakan adalah $\frac{1}{4}$ atau 0.25.

4. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan dapat diuraikan beberapa kesimpulan sebagai berikut:

1. Validitas isi soal Ujian Sekolah Berbasis Nasional Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur Tahun Pelajaran 2017/2018 belum memenuhi kriteria Validitas Isi.
2. Analisis kualitas butir soal Ujian Sekolah Berbasis Nasional Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur Tahun Pelajaran 2017/2018 berdasarkan parameter tingkat kesukaran adalah 7 (20%) butir pada kategori sangat sukar, 13 (37,14%) butir pada kategori sukar, 15 (42,86%) butir pada kategori sedang, dan tidak terdapat butir dengan kategori mudah dan sangat mudah.
3. Analisis kualitas butir Soal Ujian Sekolah Berbasis Nasional Mata

Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur Tahun Pelajaran 2017/2018 berdasarkan parameter daya beda adalah 16 (45,71%) butir pada kategori baik, 19 (54,29%) butir pada kategori buruk, dan parameter tingkat kesukaran adalah 1 (2,86%) butir pada kategori sangat sukar, 8 (22,85%) butir pada kategori sukar, 25 (71,43%) butir pada kategori sedang, 1 (2,86%) butir pada kategori mudah, dan tidak terdapat butir dengan kategori sangat mudah.

4. Analisis kualitas butir Soal Ujian Sekolah Berbasis Nasional Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur Tahun Pelajaran 2017/2018 berdasarkan parameter jawaban tebakkan adalah 20 (57,14%) butir yang berfungsi, dan 15 (42,86%) butir yang tidak berfungsi, dan parameter daya beda adalah 19 (54.29%) butir pada kategori baik, dan 16 (45,71%) butir pada kategori buruk, serta parameter tingkat kesukaran adalah 8 (22,86%) butir pada kategori sangat sukar, 9 (25,71%) butir pada kategori sukar, 10 (28,57%) butir pada kategori sedang, 8 (22.86%) butir pada kategori mudah, dan tidak terdapat butir dengan kategori sangat mudah.

5. Saran

Berdasarkan hasil penelitian dan kesimpulan yang dikemukakan dalam penelitian ini, sehubungan dengan implementasi Teori Respon Butir

dalam menganalisis kualitas butir tes USBN SMP Mata Pelajaran IPA di Kabupaten Luwu Timur maka Secara empiris, untuk mendeskripsikan kualitas tes USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur TP. 2017/2018 sebaiknya berdasarkan pada Model Logistik 3 Parameter. Hal ini didasarkan pada uji kecocokan model yang mempersyaratkan bahwa data hasil penelitian ini cocok dengan model 3 PL. Artinya bahwa model 3 PL dianggap paling maksimal memberikan informasi mengenai kualitas tes USBN Mata Pelajaran IPA SMP Negeri di Kabupaten Luwu Timur TP. 2017/2018.

6. Daftar Pustaka

- Ali, S & Khaeruddin. 2012. *Evaluasi Pembelajaran*. Makassar: Badan Penerbit UNM
- Allen, M.J & Yen, W.M. 1979. *Introduction to Measurement Theory*. California: Brooks/Cole Publishing Company.
- Arifin, Z. 2017. *Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya.
- Baker, F. B. 2011. *Dasar-Dasar Teori Pertanyaan Butir Soal*. Terjemahan oleh Bukka, M. 2011. Tidak Diterbitkan. Makassar : Jurusan Matematika UNM.
- Crocker, L & Algina, J. 2008. *Introduction to Classical and*

- Modern Test Theory*. USA: Cengage Learning.
- Daryanto. 1999. *Evaluasi Pendidikan (Komponen MKMD)*. Jakarta: PT. Rineka Cipta
- Hambleton, R.K, Swaminatan, H., & Roger, H.J. 1991. *Fundamentals of Item Response Theory*. Newbury Park, California: Sage Publications.
- Hasnah, 2015. *Analisis Kualitas Soal Matematika Ujian Sekolah Kelas XII IPA SMA Negeri di Watansoppeng Berdasarkan Teori Respon Butir*. Makassar: PEP Educational Assessment.
- Hidayati, K. 2002. *Keakuratan Hasil Analisis Butir Menurut Teori Tes Klasik Dan Teori Respon Butir Ditinjau Dari Ukuran Sampel*. Yogyakarta: Jurusan Pendidikan Matematika FMIPA UNY.
- Mansyur, Rasyid,H., Suratno. 2015. *Asesmen Pembelajaran di Sekolah*. Yogyakarta: Pustaka Pelajar
- Mardapi, D. 2017. *Pengukuran Penilaian dan Evaluasi Pendidikan*. Yogyakarta: Parama Publishing.
- Naga, D.S. 2012. *Teori Skor Pada Pengukuran Mental*. Jakarta: Nagarani Citrayasa.
- Permana W. Y. *Pengertian Pengukuran (Measurement), Penilaian (Assessment) dan Evaluasi (Evaluate) dalam Pendidikan*. (16 Oktober 2014) terdapat dalam situs <https://yogapermanawijaya.wordpress.com>. Diakses tanggal 30 Juli 2018.
- Retnawati, H. 2014. *Teori Respon Butir dan Penerapannya*. Yogyakarta: Parama Publishing.
- _____. 2016. *Validitas, Realibilitas, dan Karakter Butir*. Yogyakarta: Parama Publishing.
- Ridho, A. 2014. *Prosiding Seminar Nasional Psikometri Invariansi Sebagai Bukti Validitas Pengukuran (7-9)*. Malang: UIN Maliki Malang.
- Ruslan.2009. *Validitas Isi*. Buletin Pa'biritta. No.10. Tahun VI, 18-19.
- Salam, S., & Bangkona, D. 2012. *Pedoman Penulisan Tesis dan Disertasi Edisi Kelima*. Makassar : Kerjasama PPs UNM dan Badan Penerbit UNM.
- Santosa, A.2013. *Kamus Umum Bahasa Indonesia*. Universitas Muhammadiyah Surakarta: Mahkota Kita.
- Setyawarno, D. 2016. *Analisis Data Pengukuran Menggunakan Program Quest*. (online).(<http://www.staff.uny.ac.id> , Diakses 4 april 2019).

- Shabir, A. 2014. *Analisis Kulaitas Soal Try-Out SMP Bahasa Inggris Kota Makassar Berdasarkan Teori Respon Butir*. Tidak Diterbitkan. Makassar: Program Pascasarjana UNM.
- Sudaryono. 2011. *Implementasi Teori Respon Butir Item Response Theory Pada Penilaian Hasil Belajar Akhir Di Sekolah*. (Online), (<http://www.sudaryono.ilearning.me>), Diakses 9 Desember 2018).
- _____. 2013. *Teori Respon Butir*. Yogyakarta: Graha Ilmu.
- Sudjana, N. 2001. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya.
- Sudijono, A. 2016. *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Pers.
- Sugiono, 2017. *Metode Penelitian Kombinasi (Mixed Methods)*. Bandung: Alfabeta.
- Sukardi (2008). *Evaluasi Pendidikan Prinsip dan Operasional*. Yogyakarta: Bumi Aksara.
- Sukiman. 2012. *Pengembangan Sistem Evaluasi*. Yogyakarta: Insan Madani.
- Surapranata, S. 2006. *Analisis Validitas, Realibilitas dan Interpretasi Hasil Tes*. Bandung: Remaja Rosdakarya.
- Susetyo, B. 2015. *Prosedur Penyusunan dan Analisis Tes*. Bandung: Refika Aditama.
- Wirawan. 2016. *Evaluasi: Teori, Model, Metodologi, Standar Aplikasi dan Teori*. Jakarta: RajaGrafindo Persada.
- Yusuf, A.M. 2015. *Asesmen dan Evaluasi Pendidikan*. Jakarta: Prenadamedia Group.
- _____. *Pengertian Evaluasi Menurut Pakar*. (Tahun 2009). Terdapat dalam <https://bangfajar.wordpress.com> . Diakses tanggal 30 Juli 2018.
- _____. <https://ainamulyana.blogspot.com/2017/01/ini-dia-permendikbud-no-3-tahun-2017.html>. Diakses tanggal 5 November 2018.

