

PERBANDINGAN ESTIMASI KEMAMPUAN MATEMATIKA SISWA MENGUNAKAN PENSKORAN DIKOTOMUS DAN POLITOMUS

Abdul Manaf¹, Samritin², Surlina³

^{1,2,3}Universitas Muhammadiyah Buton

*Corresponding Author: abdulmanafumbuton@gmail.com

ABSTRAK

Penelitian ini bertujuan mengestimasi kemampuan siswa menggunakan penskoran dikotomus dan politomus berdasarkan hasil respons peserta ujian matematika SMP. Penelitian ini termasuk penelitian eksplorasi dengan pendekatan kuantitatif. Data yang dikumpulkan merupakan data sekunder. Teknik dokumentasi digunakan dalam pengumpulan data yang bersumber dari jawaban siswa dalam pelaksanaan Ujian Nasional mata pelajaran matematika di Papua. Terdapat 8002 siswa SMP yang mengikuti Ujian Nasional pada Tahun 2015. Estimasi kemampuan data dikotomus menggunakan metode MLE dengan bantuan program BILOG-MG dan estimasi kemampuan data politomus menggunakan metode EAP program PARSCALE. Hasil estimasi kemampuan siswa berdasarkan tipe penskoran data dikotomus berada pada interval nilai (1,65) sampai (-3,63), standar deviasi (1,00) dengan rerata (0,00). Hasil ini menunjukkan bahwa pada umumnya kemampuan siswa berada pada kategori sedang disekitar (0,00) dengan distribusi kemampuan normal. Sedangkan untuk data politomus nilainya berada pada interval nilai (1,96) sampai (-1,95), standar deviasi (1,00) dengan rerata (0,00). Hasil ini menunjukkan bahwa pada umumnya kemampuan siswa berada pada kategori sedang disekitar (0,00) dengan distribusi kemampuan normal. Dengan demikian akurasi estimasi kemampuan siswa SMP dengan metode MLE dan EAP tidak persis sama.

Kata Kunci : penskoran dikotomus, penskoran politomus, kemampuan siswa, matematika

ABSTRACT

This study aims to estimate students' abilities using dichotomous and polytomous scoring based on the results of the participants' responses to the junior high school math exams. This research includes exploratory research with a quantitative approach. The data collected is secondary data. Documentation techniques are used in collecting data sourced from students' answers in the implementation of the Mathematics National Examination in Papua. There were 8,002 junior high school students who took the National Examination in 2015. Estimating the capabilities of dichotomous data using the MLE method with the help of the BILOG-MG program and estimating the capabilities of polytomous data using EAP method with the help of the PARSCALE program. The results of estimating students' abilities based on the type of dichotomous data scoring are in the data interval 1.65 to -3.63, the standard deviation is 1.00 with a mean of 0.00. This shows that in general students' abilities are in the moderate category around 0 with a normal distribution of abilities. Whereas for polytomous data the value is in the value interval 1.96 to -1.95, the standard deviation is 1.00 with a mean of 0.00. This shows that in general students' abilities are in the moderate category around 0 with a normal distribution of abilities. Thus the accuracy of estimating the ability of junior high school students using the MLE and EAP methods is not exactly the same.

Keywords : dichotomous scoring, polytomous scoring, student ability, mathematics

PENDAHULUAN

Pendidikan merupakan *basic needs* dalam pembangunan manusia, sehingga pemerintah memiliki peran yang sangat urgen dalam mengelolah pendidikan nasional. Pemerintah telah menetapkan Standar Nasional Pendidikan dengan tujuan untuk memetakan mutu pendidikan, sehingga kualitas pendidikan di seluruh daerah merata. Indikator kualitas pendidikan menekankan kualitas masukan, kualitas proses dan hasil.

Banyak faktor yang mempengaruhi kualitas pendidikan di Indonesia tanpa terkecuali kualitas evaluasi pembelajaran atau hasil belajar siswa. Evaluasi pembelajaran membutuhkan metode yang tepat (Edelenbos & Van Buuren, 2005). HEFCE (2000) & Harvey (2002a) menyatakan evaluasi belajar dan mengajar mencakup isi dan penyampaian kurikulum, kinerja guru, penilaian/penilaian siswa, standar pencapaian akademik atau kompetensi.

Sebagian besar lembaga eksternal atau internal melakukan evaluasi memiliki tujuan berikut: (a) mengaudit prosedur kelembagaan, (b) mengakreditasi program studi, (c) menilai kualitas pengajaran, (d) menilai proyek/proposal penelitian, (e) menilai hasil penelitian, (f) menetapkan atau menetapkan standar akademik atau standar kompetensi, (g) memeriksa tingkat standar akademik atau standar kompetensi profesional, (h) memastikan kesesuaian dengan peraturan atau prosedur, (i) menilai kepuasan peserta dengan penyediaan layanan tingkat kelembagaan atau program, (j) mengukur keefektifan pengajaran dan pembelajaran, sistem manajemen, dan kualitas sistem (Harvey, 2002b).

Evaluasi kepada siswa difokuskan pada kemampuan siswa dalam menerima dan memahami materi yang disampaikan guru selama proses pembelajaran. Kinerja atau kemampuan siswa dalam pembelajaran dapat diketahui melalui pemberian tes (Wafa et al., 2020). Tes diklasifikasikan ke dalam empat kategori yaitu: (a) tes penempatan, (b) tes formatif, (c) tes diagnostik dan tes sumatif (Gronlund & Linn, 1990). Secara umum tes

hasil belajar yang digunakan di sekolah yaitu tes berbentuk esai dan berbentuk pilihan ganda atau majemuk.

Tes berbentuk esai dapat digunakan untuk mengukur kemampuan siswa pada level berpikir, menganalisis, mensintesis, dan mengevaluasi. (Hopkins, 1986). Tes pilihan majemuk digunakan untuk mengukur keterampilan tingkat yang lebih rendah seperti penerapan dan analisis, dan mereka dapat digunakan untuk mengukur beberapa pengetahuan objektif (Dodge, 2009).

Penskoran tes dapat menggunakan pendekatan dikotomus dan politomus. Penilaian untuk tes yang berbentuk pilihan ganda/majemuk biasanya menggunakan penskoran dikotomus dengan jawaban salah diberi skor (0) dan jawaban benar diberi skor (1). Sedangkan pada tes yang berbentuk esai biasanya menggunakan penskoran politomus. Inovasi penskoran politomus dapat dilakukan pada tes pilihan ganda (Saepuzam et al., 2021). Dengan demikian dalam penelitian ini, peneliti mencoba melakukan penskoran politomus untuk mengurangi keterbatasan model penskoran dikotomus. Keterbatasan pada aspek akurasi pengukuran, informasi diagnostik, dan ketuntasan atribut. Melakukan perubahan dari penskoran dikotomus ke penskoran politomus masih jarang dilakukan dalam mengestimasi kemampuan peserta tes. Sehingga perlu dilakukan analisis perbandingan kemampuan peserta tes yang diskor dikotomus dan diskor politomus.

Estimasi kemampuan matematika siswa dengan penskoran dikotomus dan politomus menggunakan *Item Response Theory* (IRT). Manfaat IRT meliputi analisis komprehensif dan pengurangan kesalahan pengukuran, pembuatan tes adaptif komputer, penskalaan yang berarti dari variabel laten, kalibrasi objektif dan penyetaraan, evaluasi tes dan bias item (Thomas, 2011). Lebih lanjut Thomas (2011) menyatakan Model IRT bersandar pada asumsi bahwa probabilitas peserta ujian yang menjawab benar merupakan fungsi dari dua set parameter: (a) posisinya pada variabel laten, parameter orang, dan (b) karakteristik dari item, parameter item.

Berdasarkan parameternya, IRT terdiri atas 3 model yaitu model 1 parameter logistik (tingkat kesukaran), model 2 parameter logistik (tingkat kesukaran dan daya pembeda butir), dan model 3 parameter logistik (tingkat kesukaran, daya pembeda butir, dan tebakan semu/guesing).

METODE PENELITIAN

Penelitian ini termasuk penelitian eksplorasi dengan pendekatan kuantitatif dengan tujuan untuk membandingkan estimasi kemampuan siswa menggunakan penskoran dikotomus dan politomus soal Ujian Nasional (UN) berdasarkan hasil respons siswa SMP di Papua pada Tahun 2015.

Subjek penelitian siswa SMP di Papua yang mengikuti Ujian Nasional (UN) pada matapelajaran matematika Tahun 2015. Butir-butir soal pada tes ujian nasional menjadi objek dalam penelitian ini. Tes berbentuk pilihan ganda yang terdiri atas 40 butir soal yang di skor dikotomus. Ada 18 butir soal yang di skor politomus berdasarkan 40 butir soal mengacu pada kisi-kisi Ujian Nasional.

Dalam penelitian ini, teknik dokumentasi digunakan dalam pengumpulan data yang bersumber dari jawaban siswa dalam pelaksanaan Ujian Nasioal mata pelajaran matematika di Papua. Terdapat 8002 siswa SMP yang mengikuti Ujian Nasional pada Tahun 2015.

Analisis data matematika tes UN dengan menggunakan dua pendekatan estimasi kemampuan siswa yakni menggunakan penskoran dikotomus dengan menggunakan program BILOGMG dan estimasi kemampuan siswa yakni menggunakan penskoran politomus dengan menggunakan program PARSCALE. Estimasi kemampuan siswa pada penskoran dikotomus menggunakan metode MLE dan estimasi kemampuan siswa pada penskoran politomus menggunakan EAP.

1. Estimasi Kemampuan Untuk Penskoran Dikotomus

Analisis parameter dalam IRT menggunakan model 3 parameter logistik (tingkat kesukaran, daya pembeda butir, dan

tebakan semu/guesing) dengan formula sebagai berikut (Hambelton dan Swaminathan, 1985).

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

Keterangan:

- P_i = probabilitas menjawab benar
- c = tebakan semu/guesing
- b = tingkat kesukaran butir
- a = daya pembeda butir
- D = faktor skala
- θ = skala kemampuan
- e = nilai euler

Analisis kemampuan siswa dianalisis menggunakan formula di bawah ini dengan pendekatan *Likelihood Maximum* (Retnawati, 2015).

$$\log L_i(\theta) = \left\{ \sum_{j=1}^n x_{ij} \log_e P_j(\theta) + (1 - x_{ij}) \log_e [1 - P_j(\theta)] \right\}$$

$$\frac{\partial \log L_i(\theta)}{\partial \theta} = \sum_{j=1}^n \frac{x_{ij} - P_j(\theta)}{P_j(\theta)[1 - P_j(\theta)]} \frac{\partial \log(\theta)}{\partial \theta}$$

2. Estimasi Kemampuan Untuk Penskoran Politomus

Estimasi kemapuan peserta tes untuk penskoran politomus menggunakan *Partial Credit Model* (PCM). Pengembangan lebih lanjut penskoran politomus menggunakan GPCM. Menurut Muraki & Bock (1997) dalam (Lestari 2012) bentuk atau formula matematisnya sebagai berikut.

$$P_{j|k-1,k}(\theta) = \frac{P_{jk}(\theta)}{P_{j,k-1}(\theta) + P_{jk}(\theta)} = \frac{\exp[Da_j(\theta - b_{jk})]}{1 + \exp[Da_j(\theta - b_{jk})]}$$

$$P_{jk}(\theta) = \frac{\exp \left[\sum_{v=0}^k Da_j(\theta - b_{jv}) \right]}{\sum_{c=0}^{m_j} \exp \left[\sum_{v=0}^c Da_j(\theta - b_{jv}) \right]} = \frac{\exp \left[\sum_{v=0}^k Da_j(\theta - b_j + d_v) \right]}{\sum_{c=0}^{m_j} \exp \left[\sum_{v=0}^c Da_j(\theta - b_j + d_v) \right]}$$

Muraki & Bock (1997) dalam (Lestari 2012) menyatakan bahwa peluang peserta tes

dalam setiap kategori dapat dijabarkan sebagai berikut

Kategori 0:

$$P_{j0}(\theta) = \frac{1}{1 + \sum_{c=0}^{m_j} \exp \left[\sum_{v=0}^c Da_j(\theta - b_{jv}) \right]} \cdot X_k$$

Kategori 1:

$$P_{j1}(\theta) = \frac{\exp \left[Da_j(\theta - b_{j1}) \right]}{1 + \sum_{c=0}^{m_j} \exp \left[\sum_{v=0}^c Da_j(\theta - b_{jv}) \right]}$$

Kategori k:

$$P_{jk}(\theta) = \frac{\exp \left[\sum_{v=0}^k Da_j(\theta - b_{jv}) \right]}{\sum_{c=0}^{m_j} \exp \left[\sum_{v=0}^c Da_j(\theta - b_{jv}) \right]}$$

Metode estimasi kemampuan yang digunakan yaitu *Likelihood Maximum* dan Bayes (Retnawati, 2015). Pada penelitian ini

yang digunakan untuk mengukur kemampuan siswa dengan penskoran politomus menggunakan metode *Expected a Posterior* (EAP) dengan formula matematis sebagai berikut.

$$\bar{\theta}_i = \frac{\sum_{k=1}^q X_k P(x_i | X_k) A(X_k)}{\sum_{k=1}^q P(x_i | X_k) A(X_k)}$$

HASIL DAN PEMBAHASAN

1. Karakteristik Butir

Berdasarkan butir tes Ujian Nasional yang penskorannya menggunakan dikotomus (benar = 1, salah = 0) di peroleh informasi mengenai karakteristik butirnya. Sedangkan untuk penskoran politomus dengan dilakukan tahapan pengelompokan dengan menentukan kelompok butir yang memiliki indikator soal yang sama, sehingga diperoleh 18 butir baru dari hasil pengelompokan butir UN.

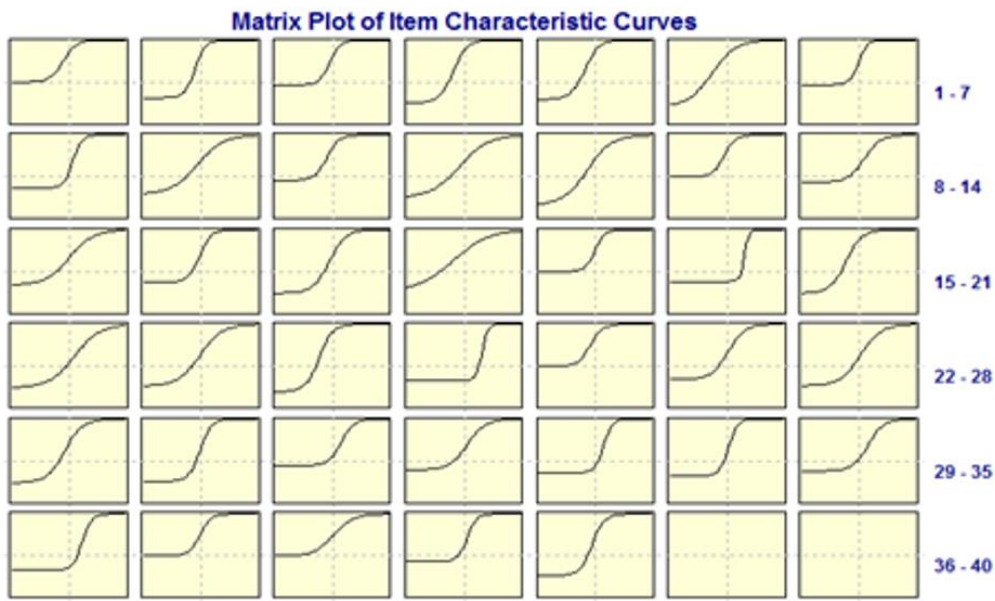
Tabel 1. Pengelompokan butir

No Item	Indikator Soal	Butir
1	Operasi hitung dan perbandingan	1 2 3
2	Bilangan berpangkat dan aritmatika sosial	4 5 6
3	Barisan bilangan dan deret	7 8
4	Pemfaktoran bentuk aljabar	9 10
5	Persamaan dan pertidaksamaan linier 1 variabel	11 12
6	Himpunan	13 14
7	Fungsi	15 16
8	Persamaan garis dan grafiknya	17 18
9	Persamaan linier dua variabel	19 20
10	Teorema pythagoras dan keliling bangun datar	21 24
11	Luas bangun datar	22 23
12	Kesebangunan dan kongruensi	25 26 27
13	Hubungan dua garis, garis istimewa pada segitiga, unsur bangun ruang	28 29 31
14	Unsur-unsur lingkaran	30
15	Kerangka dan luas permukaan bangun ruang	32 35 36
16	Volume bangun ruang	33 34
17	Ukuran pemusatan data dan penafsiran data	37 38 39
18	Peluang	40

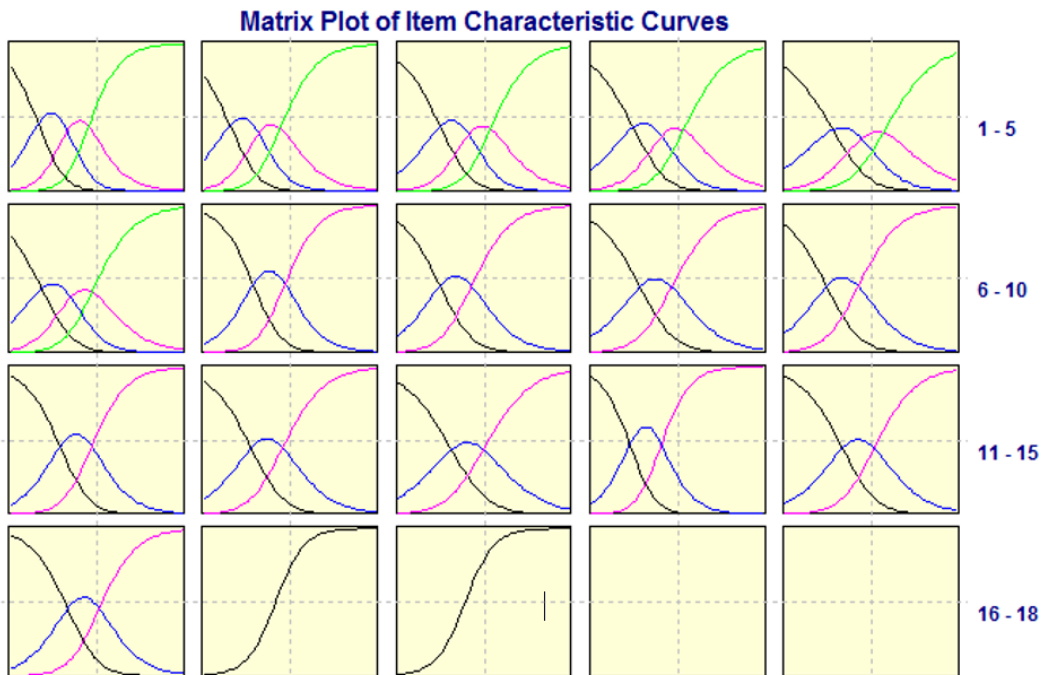
Mencermati hasil pengelompokan butir diperoleh bahwa ada 6 butir baru yang berasal dari 3 butir lama. Dengan indikator operasi hitung dan perbandingan, bilangan berpangkat dan aritmatika sosial,

kesebangunan dan Kongruensi, hubungan dua garis, garis istimewa pada segitiga, unsur bangun ruang, kerangka dan luas permukaan bangun ruang, dan ukuran pemusatan data dan penafsiran data.

Ada 10 butir baru yang masing-masing terdiri atas 2 butir lama dan 2 butir baru yang masing-masing terdiri atas 1 butir lama. Berdasarkan hasil analisis karakteristik butir untuk penskoran dikotomi diperoleh *Item Characteristik Curve* (ICC).



Gambar 1. Grafik ICC untuk penskoran dikotomus



Gambar 2. Grafik CRF untuk penskoran politomus

Berdasarkan dari Gambar 1 di atas, diketahui bahwa terdapat 40 butir soal dianalisis menggunakan IRT dengan model 3 paramter logistik. Diperoleh bahwa rerata tingkat kesukaran butir (-0,047) artinya bahwa butir dengan tingkat kesukaran sedang. Rerata daya pembeda butir (1,840) artinya bahwa butir memiliki daya pembeda yang baik. Rerata *guessing* (0,347) artinya bahwa tebakan semu tidak baik.

Berdasarkan hasil analisis karakteristik butir untuk penskoran politomus diperoleh grafik *Categorical Response Function* (CRF). Berdasarkan Gambar 2 di atas, diketahui bahwa jumlah butir yang dapat

dianalisis menggunakan metode GPCM sebanyak 18 butir. Rerata tingkat kesukaran butir (-0,774) dalam kategori sedang, rerata daya beda butir (0,862) termasuk dalam kategori baik dan rerata *guessing* (0,00).

2. Estimasi Kemampuan

Dua pendekatan estimasi kemampuan siswa yakni estimasi kemampuan siswa pada penskoran dikotomus menggunakan metode MLE dan estimasi kemampuan siswa pada penskoran politomus menggunakan EAP. Secara umum estimasi kemampuan siswa dapat disajikan sebagai berikut.

Tabel 2. Deskripsi kemampuan siswa

Tipe Data	Max	Min	Stand Dev	Rerata
Dikotomus (MLE)	1.65	-3.63	1.00	0.00
Politomus (EAP)	1.96	-1.95	1.00	0.00

Tabel 2 memberikan gambaran ukuran pemusatan kemampuan siswa berdasarkan tipe penskoran data, dari sini dapat diketahui bahwa untuk tipe data dikotomus berada pada interval data (1,65) sampai (-3,63) dengan standar deviasi (1,00) dengan rerata (0,00). Hal ini menunjukkan bahwa pada umumnya kemampuan siswa berada pada ketegori sedang disekitar (0,00) dengan sitribusi kemampuan normal. Sedangkan untuk data politomus nilainya berada pada interval nilai (1,96) sampai (-1,95) dengan standar deviasi (1,00).

Hasil penelitian ini menunjukkan bahwa akurasi estimasi kemampuan siswa SMP dengan metode MLE dan EAP tidak persis sama. Penelitian ini selaras dengan penelitian (Mahmud, 2017). Kondisi ini dipengaruhi oleh faktor-faktor dalam pengambilan sampel data (Mahmud et al., 2016). Selain itu, hasil analisis parameter butir penskoran dikotomus dan politomus ada perbedaan (Purnama, 2022), hasil penelitian Akbar (2021) diperoleh hasil estimasi kemampuan menggunakan skoring dikotomus dan politomus secara umum kemampuan siswa tidak menunjukkan perbedaan yang signifikan. Lebih lanjut

Akbar (2021) mengatakan estimasi kemampuan yang diberikan oleh EAP mencakup semua jenis pola respons juga memiliki kesalahan rata-rata yang lebih kecil dari pada pendekatanlain, termasuk MLE. Meskipun rata-rata sampel estimasi EAP tidak memberikan nilai yang bias untuk rata-rata populasi, umumnya standar deviasi sampel lebih kecil dari standar deviasi populasi. Dalam prakteknya hal ini tidak terlalu berpengaruh karena standar penyimpangan sampel akan mendekati populasi jika skornya dibakukan. Metode estimasi MLE dapat menghasilkan nilai *standar error* yang lebih kecil dibanding dengan metode estimasi EAP (Adiyo, 2014). Dengan demikian kemampuan siswa yang mengikuti Ujian Nasional pada Tahun 2015 yang dinilai dengan penskoran dikotomus dan politomus tidak hampir sama.

KESIMPULAN DAN SARAN

Hasil perbandingan estimasi kemampuan siswa SMP menggunakan penskoran dikotomus dan penskoran politomus menunjukkan bahwa akurasi estimasi kemampuan siswa dengan metode MLE dan EAP tidak persis sama. Dengan

demikian penskoran politomus dapat mengurangi keterbatasan penskoran dikotomus. Para peneliti yang ingin mengembangkan penelitian ini, sebaiknya analisis penskoran politomus dapat menggunakan *Graded Response Model* dengan bantuan R Program.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terimakasih banyak kepada semua pihak yang terlibat dalam penelitian ini. Khususnya pimpinan Universitas Muhammadiyah Buton yang telah memberikan dana penelitian sehingga penelitian ini dapat terlaksana dengan baik dan menghasilkan luaran penelitian yang bermanfaat bagi guru, kepala sekolah, pemerintah, dan akademisi/praktisi.

DAFTAR PUSTAKA

- Akbar, R. O. (2021). Comparative Analysis of Question Item Parameters and Students' Ability between Dichotomy and Polytomic Score Versions; Research on Mathematics National Exam Test Participants. *International Journal of Education and Humanities*, 1(4), 171-180.
- Adiyo, R. (2014). Pengaruh Jumlah Kategori, Rentang Threshold dan Metode Estimasi terhadap Akurasi Skor Tes pada Beberapa Model Politomi. *Jurnal Pengukuran Psikologi dan Pendidikan Indonesia*, 3(3).
- Dodge, J. (2009). *25 quick formative assessments for a differentiated classroom*. Sydney, Australia: Scholastic Teaching Resources.
- Edelenbos, J., & Van Buuren, A. (2005). The learning evaluation: a theoretical and empirical exploration. *Evaluation review*, 29(6), 591-612.
- Gronlund, N. E., & Linn, D. R. L. (1990). *Measurement and evaluation in teaching (6th ed.)*. New York, NY: Macmillan Publishing Company.
- Hambleton, R.K. & Swaminathan, H. (1985). *Item response theory*. Boston, MA: Kluwer Inc
- Harvey, L. (2002a). External quality evaluation: some international trends', keynote at the Higher Education Conference, Oslo, 22 January 2002
- Harvey, L. (2002b). Evaluation for what?. *Teaching in higher education*, 7(3), 245-263.
- Higher Education Funding Council For England (HEFCE). (2000). Consultation on employment indicator. Letter from Brian Fender, Chief Executive and Annexes A and B. (Bristol, HEFCE).
- Lestari, S. (2012). Penerapan *generalized partial credit model* dalam teori respon butir untuk menduga kemampuan Hasil tes uraian. Bogor: Institut Pertanian Bogor.
- Mahmud, J. (2017). Item Response Theory: A Basic Concept. *Educational Research and Reviews*, 12(5), 258-266.
- Mahmud J, Sutikno M, Naga DS. (2016). Variance difference between maximum likelihood estimation method and expected A posteriori estimation method viewed from number of test items. *Educ. Res. Rev.* 11(16):1579-89. doi: 10.5897/ERR2016.2807
- Purnama, D. N. (2022). Perbandingan Estimasi Kemampuan Akuntansi Siswa SMK Pada Penskoran Dikotomus Dan Politomus. *Jurnal Pendidikan Akuntansi Indonesia*, 20(2), 145-154.
- Retnawati, H. (2015). Perbandingan Estimasi Kemampuan Laten Antara Metode Maksimum Likelihood Dan Metode Bayes. *Jurnal Studi Dan Evaluasi Pendidikan*, Vol. 19 No. 2. pp. 146-155.
- Saepuzaman, D., Istiyono, E., Haryanto, H., Retnawati, H., & Yustiandi, Y. (2021). Analisis karakteristik butir soal fisika dengan pendekatan IRT penskoran dikotomus dan politomus. *Radiasi: Jurnal Berkala Pendidikan Fisika*, 14(2), 62-75.

Thomas, M. L. (2011). The value of item response theory in clinical assessment: a review. *Assessment*, 18(3), 291-307.

Wafa, M. N., Hussani, S. A. M., & Pazhman, J. (2020). Evaluation of Students' Mathematical Ability in Afghanistan's

Schools Using Cognitive Diagnosis Models. *Eurasia Journal of Mathematics, Science and Technology Education*, 16(6).

<https://doi.org/10.29333/ejmste/7834>