

ANALISIS ESTIMASI KEMAMPUAN SISWA MENJAWAB SOAL FISIKA DENGAN PENDEKATAN *ITEM RESPONSE THEORY* *PENSKORAN DIKOTOMUS DAN POLITOMUS*

Duden Saepuzaman¹, Edi Istiyono², Haryanto³, Heri Retnawati⁴, Yustiandi⁵

¹Departemen Pendidikan Fisika, FPMIPA, Universitas Pendidikan Indonesia, Bandung

^{2,3,4}Program Studi Penelitian dan Evaluasi Pendidikan, Universitas Negeri Yogyakarta, Yogyakarta

⁵SMAN CMBBS, Pandeglang

dudensaepuzaman.2019@student.uny.ac.id

Abstract: *This study compares students' abilities using the Item Response Theory (IRT) approach to dichotomous and polytomous scoring. This research is quantitative descriptive. The research subjects were 1175 high school students in class XI spread across West Java and Banten provinces, consisting of 450 male students and 725 female students. Response data with dichotomous scoring were analyzed using the item response theory approach with the BILOG-MG program, while the polytomous scoring was analyzed using the GPCM approach using the R program. The results of the model fit test showed that the items most fit with the 2PL model. The instrument used is a work and energy material test instrument tested for validity d , reliability, distinguishing power, and difficulty level. The results showed that the average ability with a polytomous score was greater than that of a dichotomous, even though compared to having almost the same ability but with a relatively different distribution. The distribution of students' abilities with polytomous scoring is closer to the normal curve than the dichotomous scoring. The relationship between students' ability scores with these two scoring approaches is shown by a correlation coefficient score of 0.990 and a determination index of 0.9808 with a prediction line $y = 0.9735x + 0.0036$.*

Keywords: *ability, item response theory, dichotomous, politomus*

Abstrak: Penelitian ini bertujuan untuk membandingkan kemampuan siswa dengan menggunakan pendekatan *Item Response Theory* (IRT) penskoran dikotomus dan politomus. Penelitian ini merupakan penelitian deskriptif kuantitatif. Subyek penelitian sebanyak 1175 siswa SMA kelas XI yang tersebar di provinsi Jawa Barat dan Banten yang terdiri 450 siswa laki-laki dan 725 siswa perempuan. Data respon dengan penskoran dikotomus dianalisis menggunakan pendekatan teori respon butir dengan program BILOG-MG sedangkan untuk penskoran politomus dianalisis dengan pendekatan GPCM menggunakan program R. Hasil uji kecocokan model menunjukkan bahwa butir soal paling banyak fit dengan model 2PL. Instrumen yang digunakan merupakan instrument tes materi usaha dan energi yang telah diuji validitas d , reliabilitas, daya beda dan tingkat kesulitannya. Hasil penelitian menunjukkan bahwa rerata kemampuan

dengan penskoran politomus lebih besar dibandingkan dikotomus, meskipun jika dibandingkan memiliki kemampuan yang hampir sama tetapi dengan sebaran yang relatif berbeda. Secara umum, sebaran kemampuan peserta didik dengan penskoran politomus lebih mendekati kurva normal dibandingkan dengan penskoran dikotomus. Terkait dengan hubungan nilai kemampuan peserta didik dengan kedua pendekatan penskoran ini ditunjukkan dengan skor koefisien korelasi sebesar 0,990 dan indeks determinasi adalah 0,9808 dengan garis prediksi $y = 0,9735x + 0,0036$.

Kata kunci: kemampuan, *item response theory*, dikotomus, politomus

Model penilaian untuk item multiple choice biasanya menggunakan model penskoran dikotomis yaitu jawaban yang benar diberi skor 1 dan jawaban salah diberi skor 0. Penggunaan penskoran ini sangat lazim digunakan karena sangat praktis, termasuk dalam penilaian hasil belajar fisika. Penerapan model penskoran dikotomus untuk menganalisis hasil tes memiliki beberapa kekurangan. Model penskoran dikotomus tersebut tidak dapat digunakan untuk membedakan kesalahan yang dilakukan oleh peserta didik, karena semua option yang salah diberi Skor 0. Padahal, kesalahan yang terjadi pada option yang salah dapat berasal dari penyebab kesalahan kategori yang berbeda. Padahal untuk menyelesaikan butir soal misalnya dalam fisika contohnya, diperlukan sejumlah tahapan penyelesaian. Pada setiap tahapan penyelesaian butir diperlukan penguasaan kompetensi yang mendasari butir soal tersebut (Gierl, 2007; Gierl, Zheng & Cui, 2008; Robert & Gierl, 2010). Hanya peserta yang sudah menguasai dan menerapkan dengan tepat yang diberi skor. Dengan kata lain, model penskoran ini belum menghargai tahapan-tahapan penyelesaian soal, karena dengan tingkat kesalahan yang berbeda mendapatkan skor yang sama yakni 0 (Istiyono, 2018). Oleh karena itu perlu inovasi model penskoran dengan mengembangkan model penskoran politomus pada butir soal objektif pilihan ganda pada Fisika.

Sebagai alternatif penskoran, respon pada tes dapat dianalisis menggunakan model penskoran politomus. Analisis respon secara politomus dapat meningkatkan akurasi pengukuran, karena itu kecenderungan

pengembangan penskoran disarankan menggunakan sistem penskoran politomus (Tognolini & Davidson, 2003). Isgiyanto (2013: 39) juga merincikan model penskoran politomus dapat memberikan solusi terhadap beberapa keterbatasan model penskoran dikotomus, terkait dengan akurasi pengukuran, ketuntasan atribut yang mendasari butir soal, dan dalam penemuan informasi diagnostik yang belum didapat dari model penskoran dikotomus.

Analisis respon yang dianalisis menggunakan penskoran dikotomus dan politomus memiliki tujuan yang sama, yaitu untuk mengukur kemampuan siswa atau peserta tes. Penelitian ini difokuskan pada estimasi kemampuan siswa menggunakan IRT penskoran dikotomus dan politomus. Kemampuan yang dimaksud dalam penelitian ini adalah kemampuan siswa dalam menjawab soal fisika yang diberikan sesuai dengan kompetensi yang dituju.

METODE

Penelitian ini merupakan penelitian deskriptif kuantitatif. Penentuan sample menggunakan *simple random sampling*. Ukuran sampel sebanyak 1177 siswa SMA kelas XI yang tersebar di provinsi Jawa Barat dan Banten yang terdiri 451 siswa laki-laki dan 726 siswa perempuan. Instrumen yang digunakan adalah instrument tes fisika materi usaha dan energi yang berjumlah 25 butir bentuk multiple choice dengan matriks indikator disajikan dalam tabel 1. Instrumen ini telah diuji validitas, reliabilitas, daya beda dan tingkat kesulitannya.

Tabel 1. Matriks Instrumen Tes Materi Usaha dan Energi

Indikator	Butir
Mendefinisikan konsep usaha	1, 2, 3
Menentukan besar usaha dengan menggunakan persamaan	4,5,6,7
Menentukan besar usaha berdasarkan grafik F-s	8,9
Menentukan besar energi potensial	10, 11, 12,13
Mendefinisikan konsep energi kinetik	14 ,15, 16
Menganalisis besar hubungan usaha dengan perubahan energy	17 ,18, 19, 20
Memahami konsep daya	21, 22
Menganalisis hukum kekekalan energy	23 , 24 , 25

Data respon dengan penskoran dikotomus dianalisis menggunakan pendekatan teori respon butir dengan program BILOG-MG sedangkan untuk penskoran politomus dianalisis dengan pendekatan GPCM menggunakan program R. Tahapan awal yang dilakukan untuk pendekatan IRT adalah menentukan kecocokan model.

Analisis untuk penskoran dikotomus menghasilkan jumlah item yang cocok dengan model 1 PL adalah 9 item, model 2 PL adalah 17 item, dan model 3 PL adalah 16 item. Jika ditinjau dari persentasenya, kesesuaian dengan model 2PL ini paling besar dibandingkan dengan 1PL dan 3 PL. Sehingga dapat disimpulkan berdasarkan analisis ini bahwa analisis instrument tes Usaha dan Energi cocok dengan model parameter 2PL. Analisis yang sama dilakukan untuk penskoran politomus yang menunjukkan bahwa data paling banyak fit dengan menggunakan model IRT penskoran politomus GPCM2PL.

HASIL DAN PEMBAHASAN

Untuk menganalisis estimasi kemampuan siswa dengan menggunakan penskoran dikotomus dan politomus, tahapan diawali

dengan membuat butir soal usaha dan energi menjadi kelompok butir sehingga respon siswa dapat diskor dengan skala dikotomus politomus. Pengelompokan ini didasarkan pada ketercapaian indikator pencapaian kompetensi untuk materi usaha dan energi. Mengacu pada pengelompokan ini diperoleh bahwa jumlah butir instrument sebanyak 25 butir untuk penskoran dikotomus, sedangkan untuk politomus sebanyak 8 butir. Artinya, soal politomus yang digunakan dalam penelitian ini merupakan soal yang sama dengan yang dikotomi, berupa multiple choice. Hanya saja dalam analisisnya tidak dipandang sebagai satu butir, tetapi dikelompokkan berdasarkan indikatornya. Artinya 1 soal dalam politomi terdiri dari beberapa butir multiple choice. Untuk analisis penskoran mengikuti banyaknya butir yang dijawab benar yang mewakili indikator tersebut. Karena tiap indikator diwakili dengan jumlah butir yang berbeda dengan mempertimbangkan kedalaman dan keluasan cakupan materi, maka jumlah butir tiap indikator (butir politomi) berbeda. Hal ini pun yang membuat banyaknya kategori skor/penilaiannya tiap butir politomi (indikator) berbeda. Secara umum pengelompokan ini disajikan dalam Tabel 2.

Tabel 2. Rincian kelompok butir untuk Penskoran Politomus

Indikator	No.Butir Dikotomus	No.Butir Politomus
1 Mendefinisikan konsep usaha	1, 2, 3	1 (4 kategori : 0,1,2,dan 3)
2 Menentukan besar usaha dengan menggunakan persamaan	4,5,6,7	2 (5 kategori : 0,1,2,3 dan 4)
3 Menentukan besar usaha berdasarkan grafik F-s	8,9	3 (3 kategori : 0,1,dan 2)
4 Menentukan besar energi potensial	10, 11, 12,13	4 (5 kategori : 0,1,2,3 dan 4)
5 Mendefinisikan konsep energi kinetik	14 ,15, 16	5 (4 kategori : 0,1,2,dan 3)
6 Menganalisis besar hubungan usaha dengan perubahan energi	17 ,18, 19, 20	6 (5 kategori : 0,1,2,3 dan 4)
7 Memahami konsep daya	21, 22	7 (3 kategori : 0,1, dan 2)
8 Menganalisis hukum kekekalan energi	23 , 24 , 25	8 (4 kategori : 0,1,2,dan 3)

Analisis menggunakan program BILOG-MG untuk penskoran dikotomus dan program R untuk penskoran politomus menghasilkan

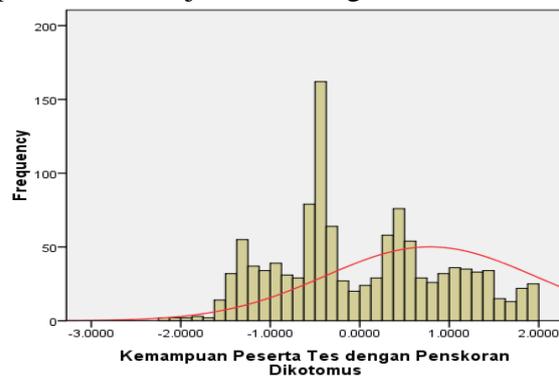
deskripsi statistik kemampuan siswa (skala -3 sampai +3) seperti disajikan pada Tabel 3.

Tabel 3. Perbandingan rerata, Standar Deviasi, Kemampuan minimum dan maksimum peserta tes dengan Penskoran Dikotomus dan Politomus

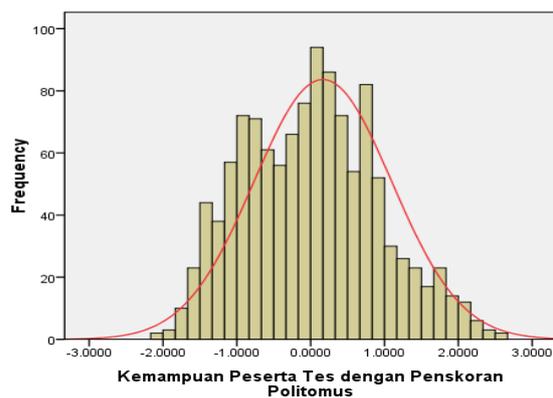
	Penskoran	
	Dikotomus	Politomus
Rerata	0,000262	0,003874
Standar Deviasi	0,9314662	0,9156666
Minimum	-2,1595	-2,1600
Maksimum	1,9944	2,6250

Tabel 3 menunjukkan bahwa ada sedikit perbedaan antara rerata, standar deviasi, kemampuan minimum dan kemampuan maksimum siswa ketika dianalisis menggunakan penskoran dikotomus dan politomus. Kemampuan merupakan kapasitas seorang individu untuk menyelesaikan tugas dalam suatu pekerjaan (Retnawati, 2016). Kemampuan yang diukur dalam penelitian ini merupakan kemampuan kognitif siswa yang berkenaan dengan hasil belajar dalam bentuk respon peserta didik pada soal materi Usaha dan Energi. Kemampuan siswa diukur berdasarkan teori respon butir dengan dua model penskoran yakni dikotomus dan politomus. Secara umum kemampuan siswa baik yang dianalisis dengan penskoran dikotomus maupun politomus menunjukkan rata-rata kemampuan yang tergolong sedang. Hal ini ditunjukkan dengan angka mean ability penskoran dikotomus sebesar 0,000262 sedangkan mean ability dengan penskoran politomus sebesar 0,003874. Hasil ini juga menunjukkan rata-rata kemampuan siswa lebih tinggi jika dianalisis dengan penskoran politomus. Selain itu, hasil analisis dengan penskoran dikotomus menunjukkan rentang kemampuan siswa secara keseluruhan berkisar antara -2,1595 hingga 1,9944. Sedangkan analisis dengan penskoran politomus menunjukkan rentang kemampuan siswa secara keseluruhan berkisar antara -2,1600 hingga 2,6250. Nilai ini semakin menegaskan bahwa kemampuan siswa dengan penskoran politomus sedikit lebih tinggi dibandingkan dengan penskoran dikotomus.

Sebaran kemampuan siswa secara keseluruhan untuk masing-masing penskoran disajikan dalam gambar 1.



(a)



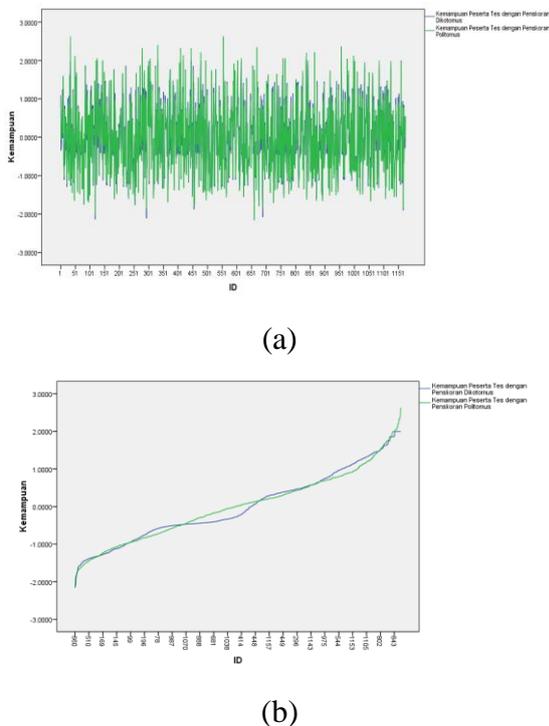
(b)

Gambar 1. Histogram sebaran kemampuan peserta tes dengan penskoran (a) dikotomus, (b) politomus

Berdasarkan Gambar 1, nampak bahwa pada analisis penskoran dikotomus, sebaran kemampuan peserta tes terbanyak dibawah

kemampuan rata-rata (di bawah teta sama dengan nol). Sedangkan pada penskoran politomus, sebaran terbanyak mendekati nilai rerata/mean yaitu disekitar nol. Jika dibuat kurva yang representative dengan sebaran ini nampak bahwa sebaran kemampuan peserta tes dengan penskoran politomus nampak mendekati kurva normal. Artinya secara umum, analisis penskoran dengan politomus lebih baik dalam memetakan kemampuan peserta tes dibandingkan dengan penskoran dikotomus. Hasil ini sejalan dengan penelitian Syam (2018) yang menunjukkan bahwa penskoran politomus dapat memberikan hasil estimasi kemampuan siswa dan karakteristik butir soal yang lebih baik dibandingkan dengan hasil estimasi dengan model penskoran dikotomus.

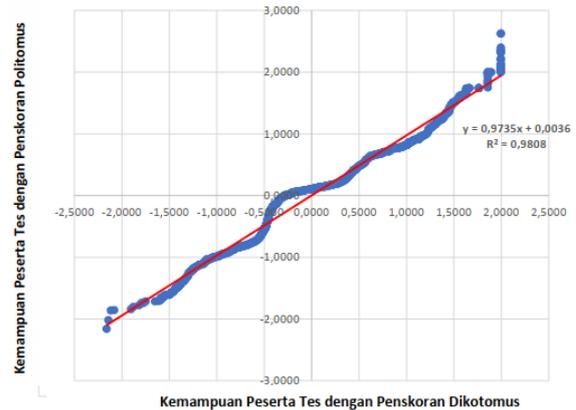
Perbedaan estimasi kemampuan peserta tes dengan dua penskoran ini dapat dilihat juga dengan melihat sebaran kemampuan setiap peserta tes untuk kedua penskoran. Secara umum untuk setiap peserta disajikan dalam gambar 2 perbedaan estimasi kemampuan ini akan lebih nampak jika kita urutkan kemampuan peserta tes mulai dari kemampuan rendah sampai kemampuan tinggi.



Gambar 2. Perbedaan estimasi kemampuan peserta tes dengan penskoran dikotomus dan

politomus (a) berdasarkan urutan ID, (b) berdasarkan urutan kemampuan

Gambar 2 menunjukkan bahwa hasil estimasi kemampuan dengan penskoran dikotomus dan politomus hampir sama atau relatif dekat untuk setiap peserta tes. Keadaan ini dapat lebih jelas dianalisis dengan menggunakan diagram petak yang disajikan pada gambar 3.



Gambar 3. Hubungan Hasil Estimasi Kemampuan Peserta Didik Menggunakan Penskoran Dikotomus dan Politomus

Gambar 3 menunjukkan bahwa adanya hubungan kemampuan siswa dengan skor dikotomus dan politomus dibuktikan dengan skor koefisien korelasi sebesar 0,990 dan indeks determinasi adalah 0,9808. Hasil lainnya menunjukkan bahwa kedua skor berkorelasi dan mendekati garis prediksi $y = 0,9735x + 0,0036$. Hasil ini sejalan dengan penelitian Syam (2018) yang meneliti tentang Estimasi Kemampuan Siswa pada Soal Matematika pada Ujian Nasional SMA Program IPA dengan Penskoran Dikotomus dan Politomus yang menunjukkan bahwa terdapat hubungan korelasi yang cukup tinggi antara kemampuan siswa yang dianalisis dengan penskoran dikotomus dan politomus.

PENUTUP

Berdasarkan hasil penelitian dapat disimpulkan bahwa pada analisis penskoran dikotomus, sebaran kemampuan peserta tes terbanyak dibawah kemampuan rata-rata (di bawah teta sama dengan nol). Sedangkan pada penskoran politomus, sebaran terbanyak

mendekati nilai rerata/mean yaitu disekitar nol dan nampak mendekati kurva normal. Artinya secara umum, sebaran kemampuan siswa dengan analisis penskoran dengan politomus lebih baik dalam memetakan kemampuan peserta tes dibandingkan dengan penskoran dikotomus.. Hasil estimasi kemampuan dengan penskoran dikotomus dan politomus hampir sama atau relatif dekat untuk setiap peserta tes. Hal ini

dibuktikan dengan skor koefisien korelasi sebesar 0,990 dan indeks determinasi adalah 0,9808. Demikian pula dengan diagram petak estimasi menggunakan penskoran dikotomus dan politomus, yang menunjukkan kedua skor berkorelasi dan mendekati garis prediksi $y = 0,9735 x + 0,0036$.

DAFTAR PUSTAKA

- Syam, Andi Sri Mardiyanti (2018) Estimasi Kemampuan Siswa dan Parameter Butir Soal Matematika pada Ujian Nasional SMA Program IPA dengan Penskoran Dikotomus dan Politomus. S2 thesis, UNY. Tidak diterbitkan, dapat diakses di : <https://eprints.uny.ac.id/view/creators/Syam=3AAndi+Sri+Mardiyanti=3A=3A.html>
- Gierl, M. J. (2007). Making diagnostic inferences about cognitive attributes using the rule-space model and attribute hierarchy method. *Journal of Educational Measurement*, 44(4), 325-340.
- Gierl, M. J., Zheng, Y., & Cui, Y. (2008). Using the attribute hierarchy method to identify and interpret cognitive skills that produce group differences. *Journal of Educational Measurement*, 45(1), 65-89.
- Isgiyanto, A. (2013). Perbandingan Penyekoran Model Rasch dan Model Partial Credit Pada Matematika. *Jurnal Kependidikan: Penelitian Inovasi Pembelajaran*, 43(1).
- Istiyono, E. (2018). Pengembangan Instrumen Penilaian dan Analisis Hasil Belajar Fisika dengan Teori Tes Klasik dan Modern. Yogyakarta : UNY Press.
- Retnawati, H. (2016). Analisis kuantitatif instrumen penelitian. *Yogyakarta: Parama Publishing*.
- Robert, M. R., & Gierl, M. J. (2010). Developing score reports for cognitive diagnostic assessment. *Educational Measurement: Issue and Practice*, 29, 25-38.
- Tognolini, J., & Davidson, M. (2003). How do we operationalise what we value? Some technical challenges in assessing higher order thinking skills. National Roundtable on Assessment Conference, Darwin, Australia.