



Analisis Karakteristik Butir Soal Fisika Dengan Pendekatan IRT Penskoran Dikotomus dan Politomus

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

¹Departemen Pendidikan Fisika, Universitas Pendidikan Indonesia

Jl. Dr. Setiabudhi No. 229, Bandung 40154, Jawa Barat, Indonesia

^{1,2,3,4}Program Studi Penelitian dan Evaluasi Pendidikan, Universitas Negeri Yogyakarta

Jl. Colombo Yogyakarta No.1, Karang Malang, Sleman, 55281, Yogyakarta, Indonesia

⁵SMA Negeri CMBBS

Jl. Raya Labuan-Pandeglang, Majasari, Pandeglang, Banten, Indonesia

dsaepuzaman@upi.edu  | DOI: <https://doi.org/10.37729/radiasi.v14i2.1200> |

Abstrak

Penelitian ini bertujuan mendeskripsikan parameter butir soal Fisika materi Usaha dan Energi 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 guna mendapatkan informasi tentang tingkat kesukaran, daya beda, kecocokan butir, estimasi kemampuan peserta didik. Hasil uji kecocokan model menunjukkan bahwa butir soal paling banyak fit dengan model 2PL. Hasil penelitian menunjukkan bahwa parameter butir soal Fisika materi Usaha dan Energi dengan penskoran dikotomus menunjukkan keseluruhan butir memiliki kriteria butir baik dan dengan penskoran politomus menunjukkan hampir seluruh butir kategori baik.

Keyword: Parameter butir, Dikotomus, Politomus

Abstract

The goal of this study is to use an item response theory (IRT) approach with dichotomous and polytomous scoring to describe the item parameters of the physics test. This is a quantitative descriptive study. 1175 high school students in class XI from West Java and Banten provinces participated in the study, comprising 450 males and 725 females. The BILOG-MG application was used to analyse response data with dichotomous scoring using the item response theory technique. In contrast, the polytomous scoring was analysed using the GPCM approach using the R program to obtain information about difficulty, discriminating power, item fit, and estimation of students' abilities. The results of the model fit test showed that most items fit the 2PL model. The study results showed that the Physics Effort and Energy items with a dichotomous scoring showed that all items had good criteria. The polytomous scoring showed almost all of the items in the good category.

Kata kunci: Items parameters, Dichotomous, Polytomous

Article Info:

Received:

09/06/2021

Revised:

19/08/2021

Accepted:

28/08/2021



1. Pendahuluan

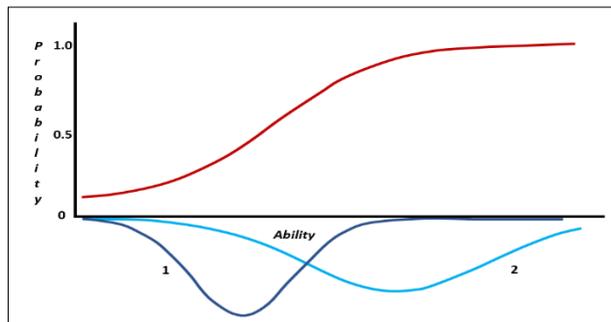
Pendidikan memiliki peranan penting dalam kemajuan sebuah bangsa dan diharapkan mampu menghasilkan generasi penerus bangsa yang unggul dan memiliki kompetensi unggul. Untuk tujuan ini, setiap negara merancang dan memberlakukan kurikulum Pendidikan yang diterapkan di sekolah, tidak terkecuali di Indonesia. Sampai saat ini pemerintah terus berupaya untuk meningkatkan kualitas pendidikan di antaranya adalah Kurikulum 2013 yang diimplementasikan secara bertahap mulai tahun pelajaran 2013/2014. Kurikulum 2013 menerapkan pembelajaran berbasis aktivitas melalui penguatan afektif, kognitif, dan keterampilan yang terintegrasi. Selain proses pembelajaran, komponen penting dalam implementasi Kurikulum 2013 adalah penilaian. Mardapi menjelaskan bahwa penilaian meliputi pengumpulan bukti-bukti tentang pencapaian belajar [1].

Dalam konteks pembelajaran, penilaian tidak bisa terpisahkan dari kegiatan proses belajar mengajar. Harlen, Gipss, Broadfoot & Nutnall [2] mengungkapkan bahwa penilaian mempunyai peran sebagai kegiatan umpan balik bagi pendidik dan peserta didik akan kemajuan belajar yang dicapai. Umpan balik ini mempengaruhi secara langsung terhadap kualitas pengalaman belajar peserta didik, apabila pendidik memanfaatkan umpan balik tersebut untuk perbaikan dalam menentukan strategi pembelajaran peserta didik. Penilaian ini sering disebut sebagai asesmen formatif atau *assessment for learning* karena bertujuan memperbaiki proses pembelajaran. Hal ini ditegaskan pada Undang-Undang No 20 tahun 2003 tentang Sistem Pendidikan Nasional pasal 58 ayat (1) yang menyatakan bahwa penilaian hasil belajar peserta didik dilakukan oleh pendidik untuk memantau proses, kemajuan, dan perbaikan hasil belajar peserta didik secara berkesinambungan.

Salah satu unsur yang harus diperhatikan dalam penilaian adalah mengusahakan dan memastikan hasil penilaian hasil belajar menggambarkan kemampuan siswa secara akurat. Suatu penilaian disebut akurat jika hasil penilaian mengandung kesalahan atau *error* sekecil mungkin. Untuk mendapatkan hasil yang akurat menggambarkan kemampuan peserta didik, kualitas instrumen tes yang digunakan harus valid, reliabel dan memiliki parameter item yang baik. Untuk tujuan ini, ada dua jenis pendekatan yang dapat digunakan untuk memperkirakan parameter item, yaitu tes klasik teori dan teori respons butir. Tes klasik teori dipandang memiliki kelemahan. Kelemahan yang paling menonjol dari teori tes klasik adalah bahwa karakteristik peserta ujian dan karakteristik tes tidak dapat dipisahkan, masing-masing dapat hanya ditafsirkan dalam konteks lain [3]. Artinya, kemampuan peserta ujian hanya ditentukan oleh tes. Ketika tes ujian akan tampak memiliki kemampuan yang lebih tinggi. Dengan kata lain, parameter item sangat tergantung pada subjek/ peserta tes begitupun sebaliknya. Karakteristik item akan berubah ketika peserta ujian berubah, dan karakteristik peserta ujian akan berubah ketika itemnya berubah. Dalam hal ini, teori tes klasik kurang dapat digunakan sebagai standar karena hasil penilaian sangat tergantung pada subyek peserta tes.

Teori respons butir merupakan solusi untuk mengatasi kelemahan yang ada pada teori tes klasik karena teori respons butir memiliki konsep melepaskan keterkaitan antara item dan sampel atau subyek peserta tes. Karakteristik/ kemampuan peserta ujian akan tetap sama meskipun mereka bekerja pada item dengan karakteristik yang berbeda, dan sebaliknya, karakteristik item akan tetap sama meskipun mereka dilakukan oleh peserta ujian dengan kemampuan yang berbeda. Selain itu juga, teori respons butir mendasarkan pada item/ butir bukan lagi pada perangkat tes. Hambleton *et al.*, [3] mengungkapkan setidaknya ada dua postulat terkait teori respons butir yaitu (a) kinerja peserta tes pada item tes dapat diprediksi (atau dijelaskan) oleh serangkaian faktor yang disebut *trait*, sifat laten, atau kemampuan; dan (b) hubungan antara kinerja atau *performance* peserta tes dan item dapat dijelaskan oleh fungsi yang meningkat secara monoton disebut dengan fungsi karakteristik item atau item *characteristic curve* (ICC). Fungsi ini menjelaskan bahwa ketika kemampuan meningkat, probabilitas responden menjawab dengan benar untuk suatu item juga meningkat.

Pada Gambar 1 dapat dilihat bahwa kelompok peserta ujian dengan kemampuan yang lebih tinggi memiliki probabilitas yang lebih besar menjawab dengan benar daripada kelompok peserta ujian dengan kemampuan rendah [4].



Gambar 1. Kurva ICC dan distribusi kemampuan dalam dua kelompok peserta tes (diadaptasi dari Hambleton *et al.*, [3])

Bentuk penilaian biasanya meliputi tes dan non tes. Untuk penilaian berupa tes umumnya bentuk tes yang digunakan adalah *multiple choice*. 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 dikotomis untuk menganalisis hasil tes memiliki beberapa kekurangan. Salah satunya bahwa model penskoran dikotomis kurang bisa untuk membedakan kesalahan yang dilakukan oleh peserta tes, karena semua *option* yang salah diberi skor nol. Padahal, kesalahan yang terjadi pada pemilihan jawaban (*option*) yang salah disebabkan pada beberapa hal yang berbeda. Selain itu juga 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 [5][6][7]. 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 [8]. 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 [9]. Isgiyanto juga merincikan model penskoran politomus dapat memberikan solusi terhadap beberapa keterbatasan model penskoran dikotomis, terkait dengan akurasi pengukuran, ketuntasan atribut yang mendasari butir soal, dan dalam penemuan informasi diagnostik yang belum didapat dari model penskoran dikotomis [10]. Penelitian ini akan difokuskan pada mendeskripsikan parameter butir soal Fisika materi Usaha dan Energi dengan menggunakan pendekatan IRT penskoran dikotomis dan politomus.

2. Metode

Penelitian deskriptif kuantitatif diterapkan dalam penelitian ini. Subyek penelitian 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 berupa instrumen tes fisika materi usaha energi berjumlah 25 *multiple choice*. Instrumen ini sudah terbukti valid dan reliabel. Valid dan reliabel merupakan acuan penting penentuan kualitas instrumen yang baik [11]. 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 guna mendapatkan informasi tentang tingkat kesukaran dan daya beda.

3. Hasil dan Pembahasan

Sebelum pada tahapan uji kecocokan model parameter yang sesuai atau *fit*, maka hal yang pertama dilakukan adalah yaitu uji dimensional apakah unidimensi atau multidimensi. Unidimensi berarti setiap item hanya mengukur satu kemampuan [12], sedangkan multidimensi berarti beberapa atau keseluruhan item mengukur lebih dari satu dimensi. Uji dimensional dalam penelitian ini dibuktikan melalui analisis faktor menggunakan SPSS. Analisis faktor dilakukan dengan terlebih dahulu melakukan analisis uji kelayakan, yaitu KMO-MSA tes dan uji Barlett. Tes KMO-MSA bertujuan untuk melihat kecukupan sampel, sedangkan Uji Barlett berfungsi untuk membuktikan homogenitas data. Analisis faktor dapat dilanjutkan jika nilai Kaiser Meyer Olkin (KMO)-MSA $> 0,5$ dan uji signifikan Barlett $< 0,05$ [13]. Berdasarkan data respon dalam studi ini, untuk penskoran dikotomus diperoleh nilai KMO-SMA dan Barlett seperti disajikan dalam Tabel 1.

Tabel 1. Uji KMO dan Bartlett's

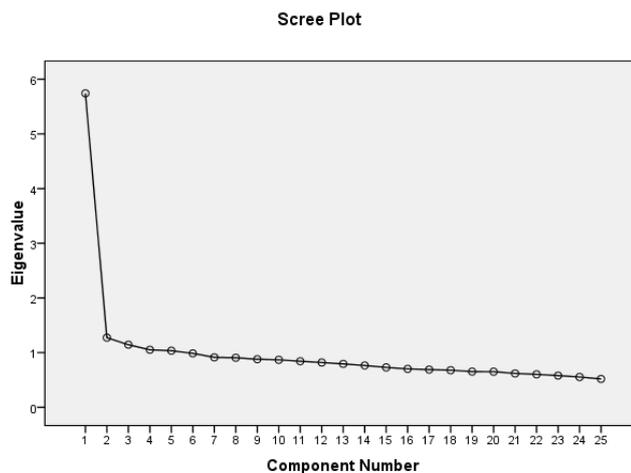
KMO dan Bartlett's		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,938
Bartlett's Test of Sphericity	Approx. Chi-Square	4889,570
	df	300
	Sig.	,000

Berdasarkan Tabel 1 dapat dilihat bahwa nilai KMO-MSA adalah 0,938 dan signifikan uji Bartlett adalah 0,000. Ini berarti sampel yang digunakan telah memenuhi persyaratan kecukupan sampel dan datanya adalah data yang homogen, sehingga analisis faktor dapat dilakukan. Hasil pengolahan data untuk analisis faktor melalui SPSS dapat dilihat di bagian nilai eigen pada Tabel 2.

Tabel 2. Nilai Eigen

Component	Initial Eigenvalues		
	Total	% of Varians	Cumulative %
1	5,742	22,969	22,969
2	1,274	5,096	28,066
3	1,145	4,578	32,644
4	1,052	4,207	36,851
...			

Berdasarkan Tabel 2, nilai eigen yang nilainya lebih dari satu menandakan satu faktor, berdasarkan nilai eigen ini, instrumen tes Usaha dan Energi memiliki tiga faktor. Tiga faktor ini dapat menjelaskan 36, 851 % varians. Nilai eigen ini kemudian dapat disajikan dalam *scree plot* pada Gambar 2.



Gambar 2. Analisis faktor *scree plot*

Scree plot dari analisis faktor menunjukkan penurunan yang sangat tajam antara faktor 1 dan faktor 2, nilai eigen kemudian mulai miring pada faktor ke-3 sehingga *scree plot* hampir membentuk sudut siku kanan. Ini menunjukkan bahwa hanya ada 1 faktor dominan dalam perangkat tes materi Usaha dan Energi.

Uji lainnya adalah independensi lokal. Dalam pembuktiannya, independensi lokal ini bisa terpenuhi jika jawaban peserta terhadap suatu butir soal tidak mempengaruhi jawaban peserta terhadap butir soal yang lain [12]. DeMars mengungkapkan bahwa independensi lokal dapat diketahui dengan membuktikan asumsi unidimensional [14]. Ini dapat diartikan bahwa jika asumsi unidimensional terpenuhi, asumsi independensi lokal juga terpenuhi. Dalam studi ini, asumsi unidimensional sudah terpenuhi sehingga uji independensi local juga sudah terpenuhi.

Untuk menentukan kesesuaian model parameter logistik yang dipakai, setidaknya ada dua cara yang bisa dilakukan [12]. Kedua cara tersebut yaitu metode statistik dan metode grafik. Metode statistik dilakukan dengan menghitung nilai *chi-square* (χ^2) dan kemudian membandingkannya dengan nilai *chi-square* (χ^2) dari tabel, atau dengan meninjau nilai probabilitas (signifikansi). Butir dikatakan cocok dengan model jika nilai *chi-square* hitung lebih kecil dari *chi-square* (χ^2) tabel atau nilai $sig > \alpha$. Untuk metode grafik dapat dilihat dari kurva karakteristik item (ICC). Melalui kurva ini, dapat dilihat seberapa tepat distribusi data dibandingkan dengan model. Model ini dikatakan cocok jika jarak titik dengan *match line* sangat dekat [12]. Penentuan akhir model parameter yang tepat atau *fit* untuk kedua model ini ditentukan dari yang paling banyak kecocokan butir dengan tipe parameter logistik (1PL, 2PL dan 3 PL) tersebut.

Dalam penelitian ini, kesesuaian model ditentukan menggunakan metode statistik yaitu dengan cara menentukan *chi-square* untuk setiap item pada setiap parameter logistik. Teknik yang digunakan pada metode ini yaitu dengan membandingkan *chi-square* hitung dengan nilai *chi-square* tabel pada derajat kebebasan tertentu. Sebuah item dipandang cocok dengan model parameter logistik jika nilai *chi-square* (χ^2) hitung tidak melebihi nilai nilai *chi-square* tabel atau kritis (χ^2_{kritis}). Kesesuaian setiap butir pada model 1 PL, 2 PL dan 3 PL disajikan pada Tabel 3.

Tabel 3. Kecocokan setiap butir pada model 1 PL , 2 PL dan 3PL

No. Butir	1PL				2PL				3PL			
	χ^2	dk	χ^2_{kritis}	Ket.	χ^2	dk	χ^2_{kritis}	Ket.	χ^2	dk	χ^2_{kritis}	Ket.
1	48,1	7	18,48	Tidak Cocok	13,1	6	16,81	cocok	28,8	7	18,48	Tidak Cocok
2	90,8	6	16,81	Tidak Cocok	19,3	7	18,48	Tidak Cocok	34,1	7	18,48	Tidak Cocok
3	116,9	8	20,09	Tidak Cocok	21,3	9	21,67	cocok	12,1	9	21,67	cocok
4	43,5	6	16,81	Tidak Cocok	4,3	6	16,81	cocok	26,3	7	18,48	Tidak Cocok
5	13,5	8	20,09	cocok	14,1	9	21,67	cocok	13,3	9	21,67	cocok
6	59,2	8	20,09	Tidak Cocok	12	8	20,09	cocok	14,4	7	18,48	cocok
7	19,6	8	20,09	cocok	5,2	9	21,67	cocok	6,2	9	21,67	cocok
8	7,8	8	20,09	cocok	7,8	9	21,67	cocok	9,4	9	21,67	cocok
9	13,8	7	18,48	cocok	6,5	8	20,09	cocok	14,7	8	20,09	cocok
10	54,5	7	18,48	Tidak Cocok	26,4	8	20,09	Tidak Cocok	30,1	8	20,09	Tidak Cocok
11	4,4	8	20,09	cocok	8	9	21,67	cocok	9,3	9	21,67	cocok
12	185,7	8	20,09	Tidak Cocok	38,1	9	21,67	Tidak Cocok	33,4	9	21,67	Tidak Cocok
13	23,1	8	20,09	Tidak Cocok	8,6	9	21,67	cocok	6,9	9	21,67	cocok
14	77	7	18,48	Tidak Cocok	38,4	9	21,67	Tidak Cocok	10,8	8	20,09	cocok
15	21,9	6	16,81	Tidak Cocok	5	7	18,48	cocok	6,7	7	18,48	cocok
16	13,4	4	13,28	Tidak Cocok	7,1	4	13,28	cocok	42,8	5	15,09	Tidak Cocok
17	6,6	7	18,48	cocok	6,3	8	20,09	cocok	27,5	8	20,09	Tidak Cocok
18	29,6	8	20,09	Tidak Cocok	23,7	9	21,67	Tidak Cocok	3,6	9	21,67	cocok
19	30,6	8	20,09	Tidak Cocok	22	9	21,67	Tidak Cocok	11,1	9	21,67	cocok
20	8,4	8	20,09	cocok	3	9	21,67	cocok	4,7	9	21,67	cocok
21	32,3	7	18,48	Tidak Cocok	4,8	8	20,09	cocok	3,1	8	20,09	cocok
22	68,4	8	20,09	Tidak Cocok	7,2	8	20,09	cocok	4,7	8	20,09	cocok
23	16,1	7	18,48	cocok	22,8	8	20,09	Tidak Cocok	29	8	20,09	Tidak Cocok
24	34,4	8	20,09	Tidak Cocok	24,3	9	21,67	Tidak Cocok	34,1	9	21,67	Tidak Cocok
25	8,9	8	20,09	cocok	16,2	9	21,67	cocok	4,7	9	21,67	cocok
Jml	cocok 1 PL			9	cocok 2 PL			17	cocok 3 PL			16

Berdasarkan Tabel 3 dapat diketahui bahwa 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 instrumen tes Usaha dan Energi cocok dengan model parameter 2PL.

Analisis lanjutan yang dilakukan yaitu mengestimasi nilai parameter item dengan merujuk model 2PL, yaitu parameter item tingkat kesulitan dan daya beda. Sebuah butir dikatakan baik jika tingkat kesulitannya (b) adalah kisaran -2 hingga +2 dan daya beda (a) ada di kisaran 0 hingga 2 [15]. Secara umum hasil estimasi parameter menggunakan model 2PL disajikan pada Tabel 4, yang mengindikasikan bahwa secara keseluruhan parameter butir tes Usaha dan Energi dengan menggunakan penskoran dikotomus berkriteria baik.

Tabel 4. Parameter butir tes materi usaha dan energi berdasarkan teori respon butir unidimensi penskoran Dikotomus model 2 parameter

Indikator	Butir	a	b	Ket
Mendefinisikan konsep usaha	1	1,154	-1,511	baik
	2	1,355	-0,981	baik
	3	0,227	1,838	baik
Menentukan besar usaha dengan menggunakan persamaan	4	1,163	-0,870	baik
	5	0,776	-0,260	baik
	6	1,075	-0,334	baik
	7	0,405	-1,119	baik
Menentukan besar usaha berdasarkan grafik F-s	8	0,690	-0,047	baik
	9	0,774	-1,170	baik
Menentukan besar energi potensial	10	1,061	-0,650	baik
	11	0,696	-0,504	baik
	12	0,170	-1,228	baik
	13	0,864	0,667	baik
Mendefinisikan konsep energi kinetik	14	0,328	-0,656	baik
	15	1,002	-0,609	baik
	16	1,123	-1,266	baik
Menganalisis besar hubungan usaha dengan perubahan energi	17	0,599	-1,425	baik
	18	0,828	0,092	baik
	19	0,864	0,476	baik
	20	0,668	0,670	baik
Memahami konsep daya	21	1,012	-0,134	baik
	22	1,096	0,179	baik
Menganalisis hukum kekekalan energi	23	0,592	-1,846	baik
	24	0,394	-1,514	baik
	25	0,598	0,129	baik

Analisis data dengan penskoran politomus disini dilakukan menggunakan analisis GPCM 2PL. Analisis ini diawali dengan membuat butir soal usaha dan energi menjadi kelompok butir [16] sehingga dapat diskor dengan skala politomus sebagaimana disajikan dalam Tabel 5.

Tabel 5. Rincian kelompok butir untuk penskoran Politomus

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

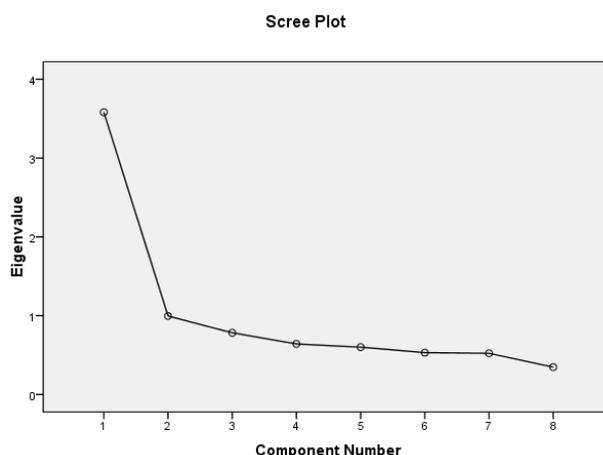
Sama halnya seperti analisis dengan penskoran dikotomus, analisis dengan politomus memerlukan asumsi. Asumsi analisis politomus GPCM, yaitu unidimensi dan independensi lokal. Pembuktiannya yang dilakukan sama seperti pada dikotomus dan terbukti unidimensi dan independensi lokal. Secara umum, hasilnya disajikan dalam Tabel 6, Tabel 7, dan Gambar 3.

Tabel 6. Uji KMO dan Bartlett's

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.868
Bartlett's Test of Sphericity	2663.149
	28
	.000

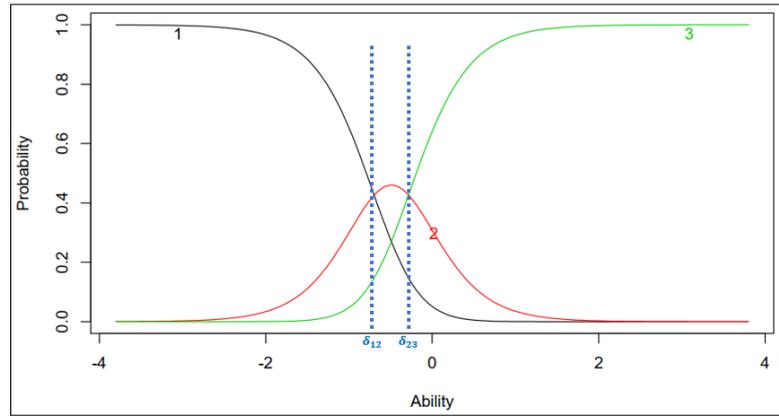
Tabel 7. Nilai Eigen

Component	Initial Eigenvalues		
	Total	% of Varians	Cumulative %
1	3.582	44.775	44.775
2	.996	12.446	57.220
3	.782	9.778	66.999
4	.641	8.014	75.013
...			



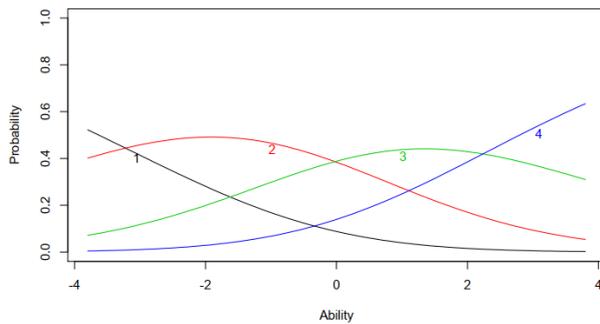
Gambar 3. Analisis Faktor *Scree plot*

Berdasarkan [Tabel 6](#) dan [Tabel 7](#) nilai eigen yang nilainya lebih dari satu menandakan satu faktor. Berdasarkan nilai eigen ini, instrumen tes Usaha dan Energi dengan penskoran politomus memiliki satu faktor. Tiga faktor ini dapat menjelaskan 44.775 % varians. Analisis *scree plot* dari analisis faktor menunjukkan penurunan yang sangat tajam antara faktor 1 dan faktor 2 yang menunjukkan bahwa hanya ada 1 faktor dominan dalam perangkat tes materi Usaha dan Energi dengan penskoran politomus. Terkait dengan parameter butir tiap item untuk penskoran dapat digambarkan dengan fungsi respon kategori (*Categorical Response Function*, CRF atau *Category Response Curve*, CRC atau *Category Characteristic Curve*, CCC) [17]. Sebagai contoh CCC untuk item nomor 7 dengan 3 kategori (kategori 1 untuk skor 0 ; kategori 2 untuk skor 1 ; kategori 3 untuk skor 2) disajikan dalam [Gambar 4](#).

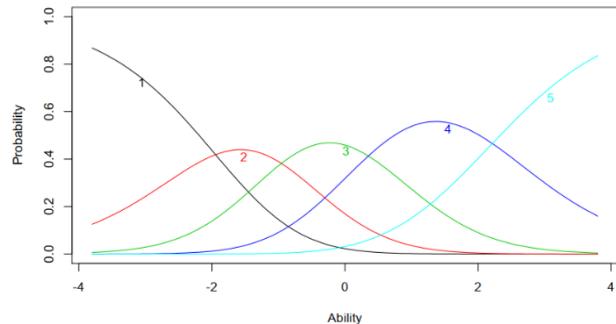


Gambar 4. CCC tiga kategori penskoran 0,1 dan 2 untuk item 7

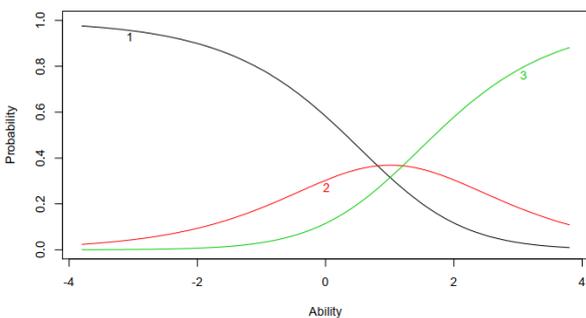
Berdasarkan **Gambar 4**, titik potong antara garis hitam, dan garis merah (1 dan 2), menyatakan peluang dan kemampuan minimal atau sering disebut juga sebagai parameter step (δ_{12}) peserta tes untuk mendapatkan skor 1. Dalam gambar nampak ada garis putus-putus biru yang menunjukkan ke nilai kemampuan $-0,698$ dengan peluang $0,4$. Ini berarti peserta tes yang mempunyai kemampuan di bawah $-0,698$ mempunyai peluang besar mendapatkan nilai 0 dengan peluang mulai dari $0,4$ sampai 1 (peluang mendapatkan nilai nol berkurang seiring besarnya kemampuan). Peserta tes yang mempunyai kemampuan $-0,698$ mempunyai peluang mendapatkan nilai 0 atau 1 sebesar $0,44$. Seseorang yang kemampuannya di atas $-0,698$ sampai $-0,284$ mempunyai peluang terbesar sebesar $0,50$ untuk mendapatkan nilai 1. Begitu pula makna yang sejalan untuk parameter step yang kedua (δ_{23}). Kurva CCC untuk keseluruhan butir politomus tes Usaha dan Energi disajikan dalam **Gambar 5**.



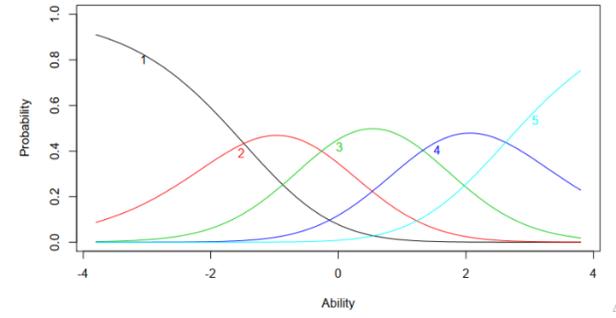
Butir 1



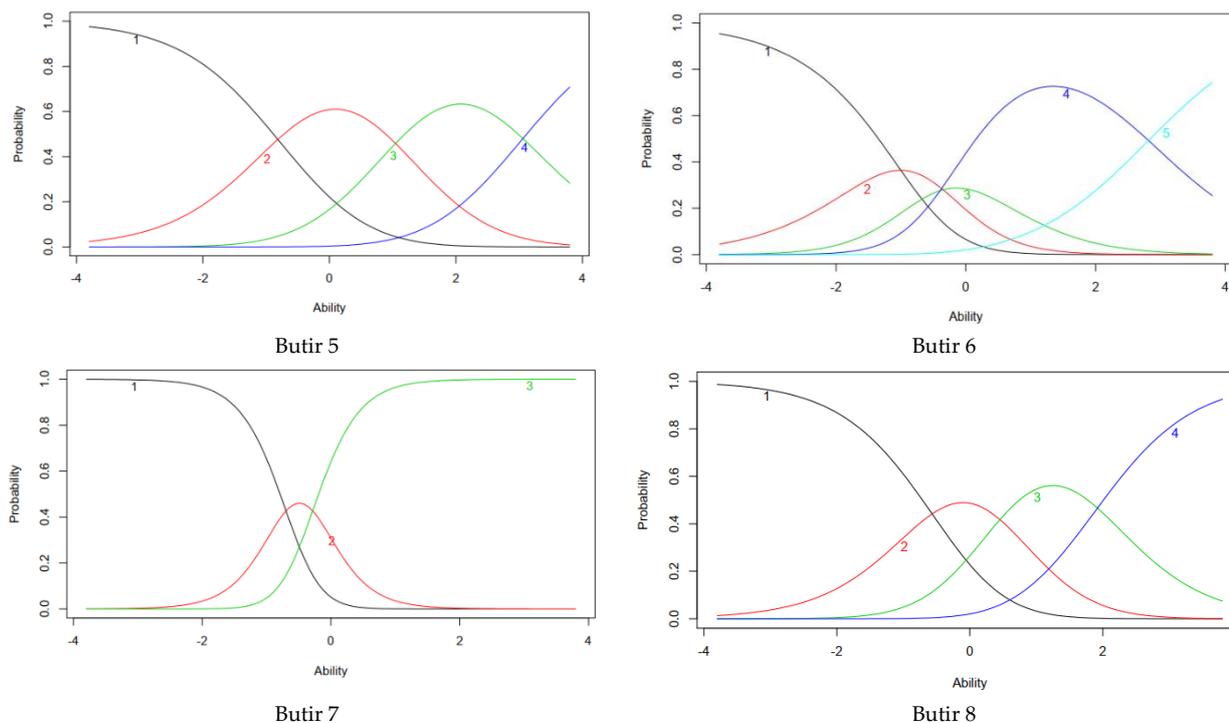
Butir 2



Butir 3



Butir 4



Gambar 5. CCC untuk setiap butir

Secara umum, parameter tiap kategori untuk item politomus disajikan dalam Tabel 8 yang menunjukkan bahwa secara keseluruhan hampir keseluruhan butir tes fisika materi usaha dan energi dengan menggunakan penskoran politomus berkriteria baik.

Tabel 8. Parameter butir tes Materi Usaha dan Energi berdasarkan teori respon butir unidimensi penskoran politomus GPCM 2 PL

Butir	Parameter kategori				Daya Beda
	1	2	3	4	
1 (4 kategori)	-3,223	-0,025	2,235		0,457
2 (5 kategori)	-1,946	-0,955	0,351	2,215	1,044
3 (4 kategori)	0,811	1,202			0,805
4 (5 kategori)	-1,481	-0,262	1,326	2,621	1,013
5 (4 kategori)	-0,816	1,045	3,060		1,242
6 (5 kategori)	-1,002	-0,350	-0,396	2,816	1,090
7 (3 kategori)	-0,698	-0,284			2,584
8 (4 kategori)	-0,566	0,453	1,920		1,340

Analisis dengan pendekatan teori respon butir penskoran dikotomus ini diawali dengan uji asumsi teori respon butir yaitu asumsi unidimensi dan independensi lokal. Asumsi unidimensi dibuktikan dengan analisis faktor eksploratori dengan menggunakan bantuan program SPSS. Analisis faktor yang dilakukan yaitu menggunakan uji Bartlett yang menghasilkan nilai KMO sebesar 0,938. Nilai KMO tersebut lebih besar dari 0,5 artinya sampel sebanyak 1175 yang digunakan dalam penelitian ini telah cukup untuk digunakan menganalisis. Berikutnya melihat nilai eigen yang berada di atas 1. Sesuai dengan **Error! Reference source not found.**, terdapat 4 faktor yang memiliki nilai eigen di atas 1, namun komponen pertama merupakan faktor dominan yang memiliki nilai eigen sebesar 5,742 dan mampu menjelaskan sebesar 24,969 total varians. Sejalan dengan hasil *scree plot* pada **Gambar 3** yang memperlihatkan adanya 1 faktor dominan yang terukur pada perangkat tes ini. Sehingga dengan demikian tes materi Usaha dan Energi memenuhi memenuhi unidimensi. Hal ini sesuai dengan pernyataan Hambleton, Swaminthan & Rogers [3] bahwa asumsi unidimensi dapat dianggap terpenuhi jika tes mengandung suatu komponen dominan yang mengukur kemampuan peserta didik. Selain itu, Wells dan Purwono [18] memberikan penegasan pada besaran presentasi varians yang ter jelaskan, jika nilainya lebih besar dari 20% maka perangkat yang diukur memuat dimensi tunggal atau bersifat unidimensi. Faktor-faktor lainnya yang juga memuat nilai eigen lebih dari 1 sangat memungkinkan untuk muncul pada tes kemampuan Fisika.

Terpenuhinya asumsi unidimensi berdasar hasil di atas menunjukkan bahwa asumsi independensi lokal juga terpenuhi [10]. Hal ini dikarenakan jika data yang dimiliki bersifat unidimensional, maka respon yang diberikan peserta tes terhadap suatu butir bersifat independen atau tidak mempengaruhi jawaban peserta tes terhadap butir soal lainnya. Artinya, kemampuan peserta tes independen terhadap butir soal Fisika materi Usaha dan Energi. Asumsi ketiga yang harus dipenuhi adalah asumsi invarian parameter butir dan kemampuan. Hasil pembuktian asumsi invarian parameter secara keseluruhan terpenuhi, baik invarian parameter butir maupun invarian kemampuan.

Analisis selanjutnya adalah uji kecocokan model. Analisis kecocokan model dilakukan secara statistik pada masing-masing model parameter logistik teori respon butir data dikotomus. Analisis kecocokan model dengan bantuan program BILOG-MG model IPL, 2PL dan 3PL secara statistik diperoleh hasil bahwa butir paling banyak cocok (*fit*) dengan model 2PL. Hal tersebut diketahui dengan membandingkan nilai signifikansi dari *chi-square*. Berdasarkan hasil pengujian kecocokan model antar model I-PL, 2-PL dan 3-PL ditemukan 9 butir cocok dengan model I-PL, 17 butir cocok dengan model 2-PL dan 16 butir cocok dengan model 3-PL; sehingga model yang digunakan dalam analisis karakteristik butir soal berdasarkan hasil analisis kecocokan model sesuai dengan banyaknya butir yang cocok pada **Tabel 3** adalah model 2-PL. Setelah menemukan model yang cocok (*fit*) yakni 2 PL maka analisis lanjutan adalah menentukan parameter butirnya seperti yang disajikan dalam **Tabel 4**. Dengan merujuk pada kriteria bahwa sebuah butir dikatakan baik jika tingkat kesulitannya (*b*) adalah kisaran -2 hingga +2 dan daya beda (*a*) ada di kisaran 0 hingga 2 diperoleh hasil bahwa secara keseluruhan parameter butir tes Usaha dan energi dengan menggunakan penskoran dikotomus berkriteria baik.

Sama halnya dengan analisis penskoran dikotomus, analisis dengan pendekatan teori respon butir penskoran politomus juga diawali dengan uji asumsi teori respon butir yaitu asumsi unidimensi dan independensi lokal. Asumsi unidimensi dibuktikan dengan analisis faktor eksploratori dengan menggunakan bantuan program SPSS. Analisis faktor yang dilakukan yaitu menggunakan uji Bartlett's yang menghasilkan nilai KMO sebesar 0,868 atau mendekati 1 dan nilai signifikansi uji Barlett sebesar $0,000 < 0,05$. Hal ini menunjukkan bahwa ukuran sampel sebesar 1175 yang diperoleh pada penelitian ini telah cukup untuk digunakan menganalisis dan tes memenuhi prasyarat analisis faktor. Berikutnya melihat nilai eigen yang berada diatas 1.

Sesuai dengan Tabel 7, terdapat 1 faktor yang memiliki nilai eigen di atas 1 yang menunjukkan komponen pertama merupakan faktor dominan yang memiliki nilai eigen sebesar 3,582 dan mampu menjelaskan sebesar 44,775 total varians. Sejalan dengan hasil *scree plot* pada Gambar 4 yang memperlihatkan adanya 1 faktor dominan yang terukur pada perangkat tes ini. Sehingga dengan demikian perangkat tes materi Usaha dan Energi memenuhi unidimensi. Terpenuhinya asumsi unidimensi berdasar hasil di atas menunjukkan bahwa asumsi independensi lokal juga terpenuhi [10]. Hal ini dikarenakan jika data yang dimiliki bersifat unidimensional, maka respon yang diberikan peserta tes terhadap suatu butir bersifat independen atau tidak mempengaruhi jawaban peserta tes terhadap butir soal lainnya. Artinya, kemampuan peserta tes independen terhadap butir soal fisika materi Usaha dan Energi dengan penskoran politomus. Analisis lanjutan juga terkait asumsi invarian parameter butir dan kemampuan terpenuhi dibuktikan dengan nilai korelasi yang positif.

Terkait analisis dalam mengestimasi parameter penskoran politomus menggunakan analisis GPCM. Hal ini sejalan dengan pendapat [19] yang menyatakan bahwa model GPCM cocok untuk menganalisis data pilihan ganda, diperkuat oleh pendapat dari Retnawati [20] yang menyatakan bahwa GPCM adalah model yang paling sesuai untuk menganalisis hasil tes dengan model penskoran politomus karena butir ini butir yang diskor dalam kategori berjenjang, namun indeks kesukaran dalam setiap langkah tidak terurut, suatu langkah dapat lebih sukar dibandingkan langkah berikutnya. Bahkan, Istiyono [7] menegaskan bahwa penggunaan GPCM untuk menganalisis tes pilihan ganda merupakan model penilaian alternatif yang adil pada pembelajaran. Estimasi parameter dengan penskoran politomus yang dihasilkan disajikan dalam Tabel 8. Dengan merujuk pada kriteria sebuah butir dikatakan baik diperoleh hasil bahwa secara keseluruhan parameter butir tes Usaha dan energi dengan menggunakan penskoran politomus berkriteria baik.

4. Kesimpulan

Berdasarkan hasil penelitian dapat disimpulkan bahwa model penskoran yang paling cocok (*fit*) dalam mengestimasi parameter butir dan kemampuan soal Fisika materi Usaha dan Energi adalah model 2-PL untuk data dikotomus dan GPCM 2PL untuk data politomus. Analisis parameter butir penskoran dikotomus menghasilkan bahwa secara keseluruhan butir dalam soal Usaha dan Energi memiliki kriteria baik dengan acuan nilai parameter butir tingkat kesulitan dan daya beda. Sedangkan dengan analisis penskoran politomus diperoleh bahwa secara keseluruhan hampir keseluruhan butir tes fisika materi usaha dan energi dengan menggunakan penskoran politomus berkriteria baik.

Daftar Pustaka

- [1] D. Mardapi, "Teknik penyusunan instrumen tes dan nontes," Yogyakarta: Mitra Cendekia, vol. 127. p. 88, 2008.
- [2] A. Pollard and J. Collins, *Reflective teaching*. A&C Black, 2005.
- [3] R. K. Hambleton, R. J. Shavelson, N. M. Webb, H. Swaminathan, and H. J. Rogers, *Fundamentals of item response theory*, vol. 2. Sage, 1991.
- [4] D. Saepuzaman, H. Haryanto, E. Istiyono, H. Retnawati, and Y. Yustiandi, "Analysis of Items Parameters on Work and Energy Subtest Using Item Response Theory," *J. Pendidik. MIPA*, vol. 22, no. 1, pp. 1–9, 2021.
- [5] M. J. Gierl, "Making diagnostic inferences about cognitive attributes using the rule-space model and attribute hierarchy method," *J. Educ. Meas.*, vol. 44, no. 4, pp. 325–340, 2007, doi: 10.1111/j.1745-3984.2007.00042.x.

- [6] M. J. Gierl, Y. Zheng, and Y. Cui, "Using the attribute hierarchy method to identify and interpret cognitive skills that produce group differences," *J. Educ. Meas.*, vol. 45, no. 1, pp. 65–89, 2008, doi: 10.1111/j.1745-3984.2007.00052.x.
- [7] M. R. Roberts and M. J. Gierl, "Developing Score Reports for Cognitive Diagnostic Assessments," *Educational Measurement: Issues and Practice*, vol. 29, no. 3, pp. 25–38, 2010, doi: 10.1111/j.1745-3992.2010.00181.x.
- [8] E. Istiyono, "Pengembangan Instrumen Penilaian dan Analisis Hasil Belajar Fisika dengan Teori Tes Klasik dan Modern." UNY Press Yogyakarta, 2018.
- [9] J. Tognolini and M. Davidson, "How do we operationalise what we value? Some technical challenges in assessing higher order thinking skills." 2003.
- [10] A. Isgiyanto, "Perbandingan Penyekoran Model Rasch dan Model Partial Credit pada Matematika," *J. Kependidikan Penelit. Inov. Pembelajaran*, vol. 43, no. 1, 2013.
- [11] D. Saepuzaman, E. Istiyono, and W. Widiastuti, "Validitas dan Reliabilitas Konstruk Instrumen Skala Sikap Siswa Terhadap Fisika Dengan Analisis Faktor Konfirmatori," *J. Ris. Pendidik. Fis.*, vol. 6, no. 1, 2021.
- [12] H. Retnawati, *Teori Respons Butir dan Penerapannya.*, 1st ed. Yogyakarta: Nuha Medika_Yogyakarta, 2014.
- [13] R. E. HAIR JR, J.F.; BLACK, W.C.; BABIN, B.J; ANDERSON, "Análise Multivariada de Dados.," Bookm. Ed., 2006.
- [14] C. DeMars, *Item response theory.* Oxford University Press, 2010.
- [15] R. K. Hambleton and H. Swaminathan, *Item response theory: principles and applications.* Boston, MA: Kluwer-Nijhoff," , 1985.
- [16] D. Sepuzaman, E. Istiyono, H. -, H. Retnawati, and Y. -, "Analisis Estimasi Kemampuan Siswa Dengan Pendekatan Item Response Theory Penskoran Dikotomus Dan Politomus," *Karst J. Pendidik. Fis. DAN Ter.*, vol. 4, no. 1, pp. 8–13, 2021, doi: 10.46918/karst.v4i1.948.
- [17] M. D. Toit, *IRT from SSI: Bilog-MG, multilog, parscale, testfact.* Scientific Software International, 2003.
- [18] H. Retnawati, *Analisis Kuantitatif Instrumen Penelitian (Panduan Peneliti, Mahasiswa, dan Psikometrian).* Yogyakarta: Parama Publishing, 2016.
- [19] D. Kaur and Y. Zhao, "Development of Physics Attitude Scale (PAS): An Instrument to Measure Students' Attitudes Toward Physics," *Asia-Pacific Educ. Res.*, vol. 26, no. 5, pp. 291–304, Oct. 2017, doi: 10.1007/s40299-017-0349-y.
- [20] H. Retnawati, S. Munadi, J. Arlinwibowo, N. F. Wulandari, and E. Sulistyaningsih, "Teachers' difficulties in implementing thematic teaching and learning in elementary schools," *New Educ. Rev.*, vol. 48, no. 2, pp. 201–212, 2017, doi: 10.15804/ner.2017.48.2.16.