



Analisis Parameter Tes Penilaian Akhir Semester Fisika Kelas X Dengan Teori Respon Butir

Maulina Nurriska Setiawati^{1*}, Purwanto², Taufik Ramlan Ramalis³

^{1,2,3}Departemen Pendidikan Fisika, Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam

Universitas Pendidikan Indonesia

*maulinanurriska99@gmail.com

Article Info

Received:
16/02/2022

Revised:
25/02/2022

Accepted:
28/02/2022

Abstrak

Hasil studi pendahuluan yang dilakukan melalui wawancara terhadap beberapa guru fisika di salah satu sekolah menengah atas Kota Bandung menunjukkan bahwa analisis kualitas dari instrumen tes yang digunakan dilakukan jika dibutuhkan saja dengan kata lain analisis instrumen tes tidak dilakukan secara berkala sehingga tidak dapat mengetahui bagaimana perkembangan kualitas dari instrumen tes yang diberikan kepada peserta didik. Meskipun dilakukan analisis, analisis yang digunakan adalah analisis klasik. Kualitas dari instrumen tes yang digunakan dapat diketahui melalui analisis parameter tes. Oleh karena itu, dilakukan penelitian yang bertujuan untuk menganalisis parameter tes penilaian akhir semester mata pelajaran Fisika kelas X tahun ajaran 2020/2021 dengan teori respon butir. Metode yang digunakan pada penelitian ini adalah metode deskriptif kuantitatif. Partisipan dalam penelitian ini adalah 215 peserta didik kelas X MIPA di salah satu SMA Kota Bandung. Seluruh data yang diperoleh dianalisis menggunakan teori respon butir model 2 PL. Model ini dipilih karena memiliki nilai puncak informasi tertinggi, yaitu 25,132. Hasil analisis menunjukkan bahwa soal penilaian akhir semester ini memiliki 2 butir soal yang tidak valid dari 25 butir soal pilihan ganda (butir ke-5 dan ke-12), nilai parameter daya pembeda (a) sebesar 0,354 yang dikategorikan baik, dan nilai parameter tingkat kesukaran (b) sebesar -1,4 yang berada dalam kategori mudah. Soal penilaian akhir semester ini memiliki nilai puncak fungsi informasi sebesar 25,132 dan SEM (*Standard Error of Measurement*) sebesar 0,199 dengan perpotongan antara kedua kurva tersebut berada pada rentang -3,619 sampai 1,206 sehingga soal tersebut reliabel untuk mengukur kemampuan sangat rendah hingga tinggi.

Kata kunci: Analisis Parameter Tes, Penilaian Akhir Semester, Teori Respon Butir

Abstract

The results of a preliminary study conducted through interviews with several physics teachers at one of the high schools in Bandung showed that the quality of the test instrument was analysed only if needed or not carried out regularly so we couldn't know how the quality of the test instruments. Although the analysis was carried out, the analysis used classical analysis. The quality of the test instrument can be seen from the characteristics of the test itself which is obtained by analyzing the parameters items on the instrument. Therefore, a study was conducted that aims to analyse the parameters of the final semester assessment test for Physics class X for the 2020/2021 academic year with item response theory. The method used in this research is descriptive quantitative method. The participants in this study were 215 students of class X MIPA in one of the high schools in Bandung. All of the data were then analysed using the item response

theory of the 2 PL model. This model was chosen because it has the highest peak value of information, which is 25.132. The results of the analysis show that the Final Semester Assessment test have 2 items that are not valid from 25 items (5th and 12th items), the value of the discriminating power parameter (a) is 0.354 which is categorized as good, and the value of the level of difficulty parameter (a) is 0.354. b) of -1.4 which is in the easy category. This Final Semester Assessment test has a peak value of the information function of 25.132 and SEM (Standard Error of Measurement) of 0.199 with the intersection between the two curves in the range of -3.619 to 1.206 so that the question is reliable to measure very low to high abilities.

Keyword: Parameter Test Analysis, Final Semester Assessment, Item Response Theory



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

Pendahuluan

Dalam dunia pendidikan, penilaian merupakan hal yang mutlak untuk dilakukan terutama penilaian hasil belajar peserta didik. Hal ini dikarenakan menurut Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 23 Tahun 2016 Tentang Standar Penilaian Pendidikan, tujuan dari penilaian hasil belajar peserta didik adalah untuk memantau dan mengevaluasi proses belajar, kemajuan belajar, dan perbaikan hasil belajar peserta didik secara berkesinambungan serta untuk menilai pencapaian Standar Kompetensi Kelulusan semua mata pelajaran di satuan pendidikan maupun nasional. Melalui penilaian hasil belajar, dapat diketahui juga seberapa jauh kemampuan peserta didik dalam menguasai bahan ajar yang telah disampaikan selama proses pembelajaran sehingga dapat berfungsi sebagai umpan balik untuk perbaikan proses belajar mengajar. Oleh sebab itu, penilaian hasil belajar peserta didik menjadi hal yang penting untuk dilaksanakan di sekolah guna mengetahui keberhasilan dari sebuah proses pembelajaran.

Penilaian hasil belajar dapat dilakukan melalui dua metode yaitu metode tes dan metode non tes (Ratnawulan dan Rusdiana, 2014). Penilaian Akhir Semester (PAS) merupakan salah satu kegiatan penilaian hasil belajar dengan metode tes. Salah satu komponen wajib dalam penilaian hasil belajar adalah instrumen penilaian yang digunakan. Hal ini dikarenakan melalui instrumen penilaian yang digunakan, efektivitas dari penilaian hasil belajar dalam mengukur pencapaian kompetensi peserta didik dapat diketahui. Oleh sebab itu, instrumen Penilaian Akhir Semester yang digunakan harus memiliki kualitas yang baik agar dapat secara efektif mengukur tingkat pencapaian kompetensi peserta didik. Kualitas instrumen Penilaian Akhir Semester dapat dilihat dari karakteristik tes itu sendiri yang didapatkan dengan cara menganalisis butir soal pada instrumen tersebut (Sopiah, Sidauruk, dan Asi, 2019). Akan tetapi, permasalahan umum yang terjadi di sekolah adalah instrumen penilaian yang digunakan jarang dilakukan analisis sehingga tidak dapat diketahui kualitas dari instrumen yang digunakan dan jika dilakukan analisis instrumen, analisis yang dilakukan menggunakan teori tes klasik (Astuti, 2019).

Namun, teori tes klasik memiliki beberapa keterbatasan. Pertama, teori tes klasik *classical test theory/CTT*) bergantung kepada sampel yang dimana estimasi parameter soal bergantung pada karakteristik peserta tes. Hal ini berpengaruh pada tingkat kesukaran soal. Tingkat kesukaran soal akan menjadi rendah apabila tes diberikan kepada peserta didik berkemampuan tinggi dan sebaliknya tingkat kesukaran soal akan tinggi apabila tes diberikan kepada peserta didik berkemampuan rendah. Kedua, kesalahan pengukuran (*standar error*) dianggap sama untuk semua peserta didik yang mengikuti tes (Azevedo, Oliveira, dan Beites, 2019). Selain itu, keterbatasan dari teori tes klasik menurut Saifuddin (dalam Sarea dan Ruslan, 2019) adalah estimasi kemampuan peserta didik bergantung pada karakteristik butir soal. Estimasi kemampuan peserta didik akan rendah apabila soal yang diberikan berada di atas kemampuannya dan sebaliknya estimasi kemampuan peserta didik akan tinggi apabila soal yang diberikan berada di bawah kemampuannya. Oleh sebab itu, munculah sebuah teori tes modern yang disebut dengan teori respon butir (*item respon theory / IRT*) untuk melengkapi keterbatasan dalam pengukuran klasik yang dimana dengan menggunakan teori ini sebuah tes dinilai dan dianalisis berdasarkan masing-masing item pada tes tersebut sehingga melalui item ini kemampuan peserta didik dapat diperhitungkan.

Teori respon butir mulai dikembangkan pada tahun 1950-an yang didasarkan pada tiga asumsi, yaitu 1) unidimensi, setiap butir tes hanya mengukur satu kemampuan laten saja, 2) independensi lokal, jawaban peserta tes terhadap suatu butir soal tidak mempengaruhi jawaban peserta tes pada butir soal lainnya, dan 3) Invariansi parameter, karakteristik butir soal tidak bergantung pada distribusi parameter kemampuan peserta tes dan parameter yang menjadi ciri dari peserta tes tidak bergantung dari ciri butir soal (Eaton, Johnson, Frank, dan Willoughby, 2019).

Dalam teori respon butir, terdapat 3 model yang dapat digunakan untuk melakukan analisis tes berbentuk dikotomi, yakni model 1 parameter logistik yang menggunakan satu parameter butir soal, yaitu tingkat kesukaran (b_i), model 2 parameter logistik yang menggunakan dua parameter butir soal, yaitu daya pembeda (a_i) dan tingkat kesukaran (b_i), dan model 3 parameter logistik yang menggunakan tiga parameter butir soal, yaitu daya pembeda (a_i), tingkat kesukaran (b_i), dan faktor tebakan (c_i) (Hambleton, Swaminathan, dan Rogers., 1991). Untuk mengetahui model parameter logistik yang sesuai dengan data yang dianalisis, dapat dilakukan dengan cara melihat puncak fungsi informasi tertinggi pada kurva fungsi informasi total di setiap model parameter logistik. Dalam penelitian ini, model parameter logistik yang sesuai dengan soal penilaian akhir semester fisika kelas X adalah model 2 parameter logistik. Secara matematis, model 2 parameter logistik dirumuskan sebagai berikut.

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

Keterangan:

- $P_i(\theta)$: probabilitas peserta tes dengan kemampuan θ dapat menjawab butir ke- i dengan benar
 θ : tingkat kemampuan peserta tes
 e : bilangan natural yang nilainya mendekati 2,178
 i : 1, 2, 3, ..., n
 n : banyaknya butir dalam tes
 b_i : indeks tingkat kesukaran butir ke- i
 a_i : indeks daya beda butir ke- i

Kemudian, untuk menguji validitas empirik pada soal penilaian akhir semester yang digunakan dapat ditentukan dari kecocokan setiap butir soal terhadap model parameter logistik yang digunakan. Kecocokan antara butir soal dengan model parameter logistik yang digunakan dapat dilihat melalui nilai *chi-square* atau *P-values* yang terdapat pada hasil keluaran aplikasi e-IRT yang ditandai dengan tulisan berwarna merah apabila butir tersebut tidak cocok dengan model parameter logistik yang digunakan. Dikatakan cocok atau valid apabila χ_{hitung}^2 (Chi-square hitung) $\leq \chi_{tabel}^2$ atau P-Value $\geq 0,05$.

Selain itu, untuk menjelaskan kekuatan suatu butir pada sebuah perangkat tes dapat diketahui melalui fungsi informasi. Melalui fungsi informasi juga, dapat diketahui soal penilaian akhir semester ini akan reliabel jika diujikan pada peserta didik dengan kemampuan dalam kategori tertentu. Secara sistematis, fungsi informasi dapat dituliskan sebagai berikut.

$$I_i(\theta) = \frac{[P_i'(\theta)]^2}{P_i(\theta)Q_i(\theta)} \quad (1)$$

Keterangan:

- $I_i(\theta)$: fungsi informasi butir ke- i
 $P_i'(\theta)$: turunan pertama fungsi $P_i(\theta)$ terhadap θ
 $P_i(\theta)$: probabilitas peserta tes dengan kemampuan θ dapat menjawab butir ke- i dengan benar
 $Q_i(\theta)$: probabilitas peserta tes dengan kemampuan θ dapat menjawab butir ke- i dengan salah ($1-P_i(\theta)$)

Hambleton dan Swaminathan (dalam Retnawati, 2014) mengungkapkan bahwa fungsi informasi tes merupakan jumlah dari fungsi informasi butir penyusun tes yang secara matematis dapat dirumuskan sebagai berikut.

$$I_i(\theta) = \sum_{i=1}^n I_i(\theta) \quad (2)$$

Hasil estimasi butir tidak terlepas dari kesalahan pengukuran. Dalam teori respon butir, kesalahan pengukuran ini dinamakan *SEM (Standard Error of Measurement)* yang dimana SEM ini memiliki hubungan dengan fungsi informasi yaitu berbanding terbalik kuadratik. Semakin besar nilai fungsi informasi butir maka semakin kecil nilai SEM nya. Semakin kecil nilai SEM nya, maka semakin tepat dan reliabel pengukuran yang dilakukan (Retnawati, 2014). Secara matematis, SEM dapat dituliskan sebagai berikut.

$$SEM(\hat{\theta}) = \frac{1}{\sqrt{I(\hat{\theta})}} \quad (3)$$

Berdasarkan uraian di atas, maka dilakukan penelitian dengan judul “ Analisis Parameter Tes Penilaian Akhir Semester Fisika Kelas X Dengan Teori Respon Butir”.

Metode

Metode penelitian yang digunakan adalah metode deskriptif kuantitatif. Penelitian deskriptif menggambarkan atau mendeskripsikan keadaan objek yang diamati secara apa adanya yang dimana pada penelitian ini objek yang akan dideskripsikan adalah parameter tes. Kemudian, pendekatan yang digunakan adalah pendekatan kuantitatif karena data yang telah diperoleh akan diwujudkan dalam bentuk angka yang dianalisis secara statistik menggunakan teori respon butir dengan bantuan program e-IRT.

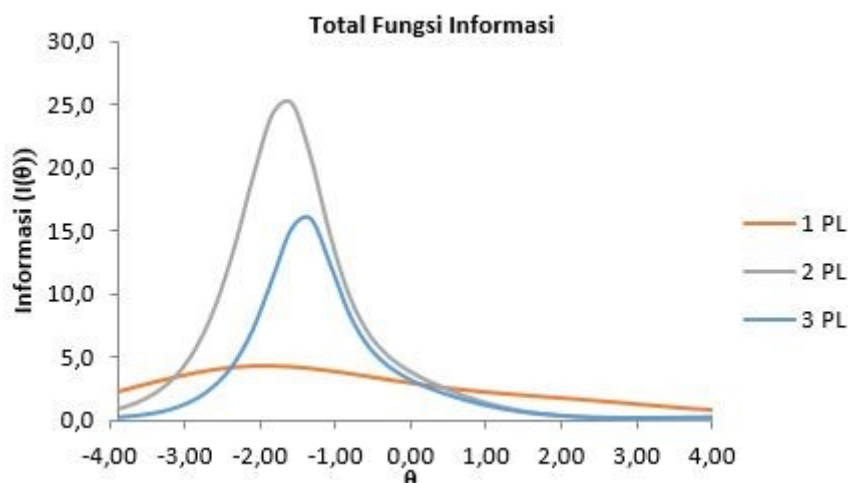
Populasi dalam penelitian ini adalah seluruh peserta didik kelas X yang telah mengikuti penilaian akhir semester tahun ajaran 2020/2021 di salah satu SMA Negeri Kota Bandung. Sampel dalam penelitian ini adalah seluruh peserta didik kelas X MIPA yang telah mengikuti penilaian akhir semester mata pelajaran Fisika tahun ajaran 2020/2021 di salah satu SMA Negeri Kota Bandung. Jumlah sampel pada penelitian ini adalah sebanyak 215 peserta didik dan dipilih dengan teknik *purposive sampling*.

Instrumen penilaian akhir semester mata pelajaran Fisika tahun ajaran 2020/2021 disusun oleh tim guru Fisika di salah satu SMA negeri Kota Bandung yang menjadi tempat penelitian. Instrumen penilaian akhir semester ini diujikan kepada 215 peserta didik kelas X MIPA di SMA tersebut oleh pihak sekolah. Hasil dari uji instrumen tersebut kemudian dimasukkan ke dalam *Microsoft Excel* oleh pihak IT sekolah tersebut. Setelah mendapatkan data hasil uji instrumen ini dari pihak sekolah, peneliti kemudian menyeleksi kembali data tersebut untuk mendapatkan data yang dibutuhkan. Kemudian, diolah menggunakan aplikasi e-IRT yang terhubung dengan *Microsoft Excel*. Analisis parameter tes penilaian akhir semester fisika kelas X dilakukan berdasarkan model 2 parameter logistik yang dilakukan dengan menggunakan bantuan program e-IRT.

Hasil dan Pembahasan

Pemilihan Model Parameter Logistik

Untuk mengetahui model parameter logistik yang sesuai dengan soal penilaian akhir semester fisika kelas X, dapat dilakukan dengan cara melihat puncak fungsi informasi tertinggi pada kurva fungsi informasi total di setiap model parameter logistik. Berikut ini merupakan grafik fungsi informasi total dari model 1PL, 2 PL, dan 3 PL yang disajikan pada gambar 1.



Gambar 1. Grafik fungsi informasi total model 1 PL, 2 PL, dan 3 PL

Berdasarkan gambar grafik fungsi informasi di atas, dapat diketahui bahwa model yang memiliki puncak grafik fungsi informasi tertinggi adalah model 2 PL dengan nilai informasi sebesar 25,132. Nilai informasi pada puncak grafik fungsi informasi 1 PL adalah 4,342 dan nilai informasi pada puncak grafik fungsi informasi model 3 PL adalah 16,014. Jadi, dapat disimpulkan bahwa model parameter logistik yang sesuai dengan soal penilaian akhir semester mata pelajaran Fisika kelas X adalah model 2 parameter logistik (2 PL).

Validitas Tes

Validitas yang digunakan adalah validitas empirik. Pada teori respon butir, validitas empirik adalah menentukan butir soal mana yang sesuai atau cocok dengan model parameter logistik yang digunakan yang dapat dilihat melalui nilai *chi-square* atau *P-values* yang terdapat pada hasil keluaran aplikasi e-IRT. Berikut ini hasil keluaran e-IRT terkait nilai *chi-square* dan *P-Values* setiap butir soal penilaian akhir semester fisika kelas X terhadap model 2 parameter logistik yang disajikan pada tabel 1.

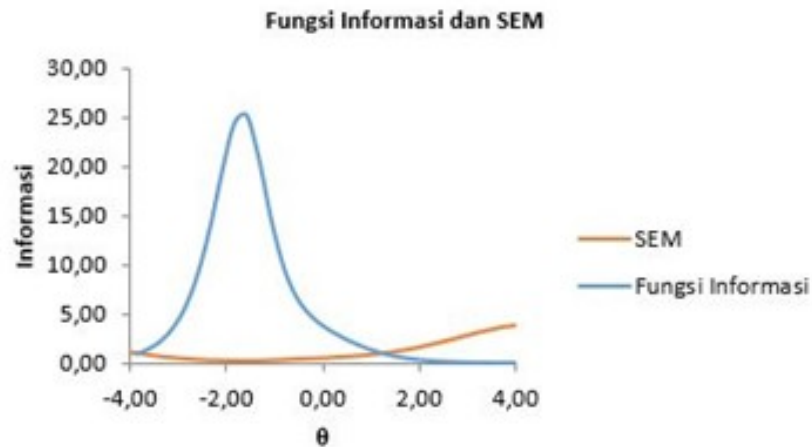
Tabel 1. Nilai *chi-square* dan *P-Values*

Item	Chi-square	Degrees of freedom	P-Value	Keterangan
1	0,425	9	1,000	Cocok
2	2,074	9	0,990	Cocok
3	9,275	9	0,412	Cocok
4	0,556	9	1,000	Cocok
5	149,460	9	0,000	Tidak Cocok
6	1,466	9	0,997	Cocok
7	0,662	9	1,000	Cocok
8	1,610	9	0,996	Cocok
9	2,322	9	0,985	Cocok
10	1,849	9	0,994	Cocok
11	5,228	9	0,814	Cocok
12	103,530	9	0,000	Tidak Cocok
13	1,101	9	0,999	Cocok
14	0,970	9	1,000	Cocok
15	10,226	9	0,333	Cocok
16	0,502	9	1,000	Cocok
17	0,378	9	1,000	Cocok
18	1,111	9	0,999	Cocok
19	0,172	9	1,000	Cocok
20	1,022	9	0,999	Cocok
21	1,862	9	0,993	Cocok
22	2,335	9	0,985	Cocok
23	1,859	9	0,994	Cocok
24	0,615	9	1,000	Cocok
25	4,859	9	0,846	Cocok

Berdasarkan tabel 1, dapat diketahui bahwa terdapat 2 butir soal penilaian akhir semester fisika kelas X yang tidak cocok dengan model 2 parameter logistik, yaitu butir soal ke-5 dan ke-12. Hal ini dikarenakan nilai χ^2_{hitung} (Chi-square hitung) $>$ χ^2_{tabel} yang dimana nilai χ^2_{tabel} adalah sebesar 16,9 dengan $\alpha = 5\%$ atau P-Value $<$ 0,05. Nilai χ^2_{hitung} untuk butir ke-5 adalah sebesar 149,460 dan P-Value butir ke-5 adalah sebesar 0,000. Nilai χ^2_{hitung} untuk butir ke-12 adalah sebesar 103,530 dan P-Value butir ke-12 adalah sebesar 0,000. Oleh sebab itu, dapat disimpulkan bahwa terdapat 2 butir soal penilaian akhir semester fisika kelas X yang tidak valid dan terdapat 23 butir soal yang valid.

Reliabilitas Tes

Reliabilitas tes atau soal dalam teori respon butri dapat ditunjukkan melalui fungsi informasi. Kebenaran dari fungsi informasi bersifat probabilitas dan tidak terlepas dari kesalahan pengukuran atau *Standart Error of Meassurement* (SEM). Oleh karena itu, perlu diketahui nilai kesalahan pengukuran atau SEM dari soal penilaian akhir semester Fisika kelas X. Kurva fungsi informasi dan SEM untuk model 2 PL disajikan pada gambar 2.

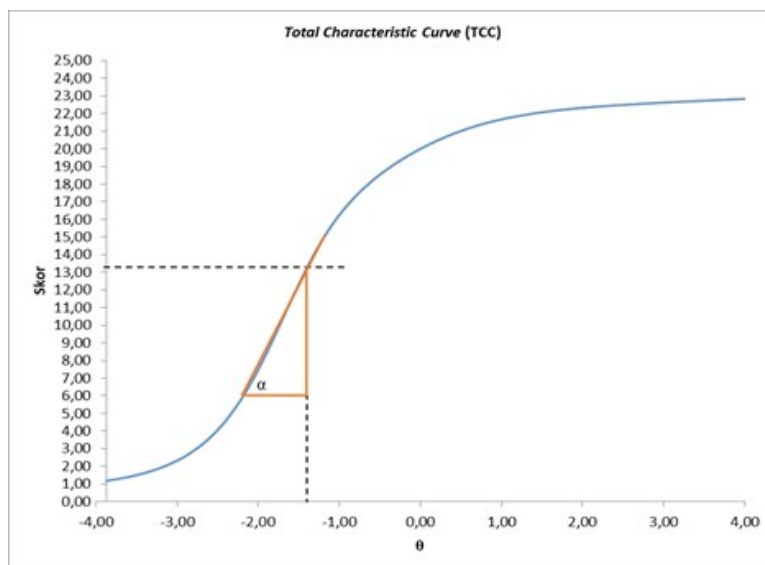


Gambar 2. Kurva fungsi informasi dan SEM model 2 PL

Berdasarkan kurva di atas, dapat diketahui bahwa puncak informasi total model 2 PL berada pada nilai informasi sebesar 25,132 dengan taraf kemampuan (θ) sebesar -1,587 dan SEM sebesar 0,199. Perpotongan antara kurva fungsi informasi dengan kurva SEM berada pada rentang -3,619 sampai dengan 1,206. Hal ini menunjukkan bahwa butir-butir soal penilaian akhir semester Fisika kelas X reliabel untuk diujikan pada peserta didik yang memiliki kemampuan dalam rentang -3,619 sampai 1,206 yaitu peserta didik dengan kategori kemampuan sangat rendah hingga kemampuan tinggi.

Parameter Tes

Berdasarkan teori respon butir, parameter tes soal secara keseluruhan dapat diketahui dengan melakukan analisis terhadap kurva karakteristik total atau *Total Characteristic Curve* (TCC). Kurva karakteristik total ini akan menunjukkan hasil estimasi skor yang diperoleh peserta didik dalam mengerjakan 25 butir soal penilaian akhir semester fisika kelas X untuk setiap tingkat kemampuan. Kurva karakteristik total model 2 PL disajikan pada gambar 3.



Gambar 3. Kurva karakteristik total (TCC) 2 PL

Berdasarkan gambar 3, diketahui bahwa peserta didik dengan kemampuan $\theta = -3,87$ (kemampuan paling rendah) memperoleh skor 1 dari skor total 25. Artinya, peserta didik dengan tingkat kemampuan tersebut hanya mampu menjawab benar 1 butir soal dari jumlah total 25 butir soal. Sedangkan, peserta didik dengan kemampuan $\theta = 4$ (kemampuan paling tinggi) memperoleh skor 23 dari skor total 25. Artinya, peserta didik dengan tingkat kemampuan tersebut akan mampu menjawab benar 23 butir soal dari jumlah total 25 butir soal. Dengan begitu, dapat disimpulkan bahwa dalam rentang kemampuan θ sama dengan $-3,87$ sampai dengan 4, rentang skor yang diperoleh yaitu 1 sampai dengan 23.

Probabilitas 1 berada pada skor 25, artinya kemampuan $\theta = -3,87$ (kemampuan paling rendah) akan memperoleh probabilitas 0,047 dari probabilitas maksimal 1 dan kemampuan $\theta = 4$ (kemampuan paling tinggi) akan memperoleh probabilitas 0,913 dari probabilitas maksimal 1. Dengan begitu, dapat diketahui bahwa probabilitas 0,5 berada pada 0,53 yaitu pada skor 13,08.

Nilai parameter b (tingkat kesukaran) soal dapat diketahui dengan cara menarik garis horizontal pada nilai probabilitas 0,5 (pada skor 13,08) sampai pada kurva TCC. Kemudian, dari perpotongan kedua garis tersebut ditarik garis secara vertikal sampai ke sumbu kemampuan (θ). Nilai yang ditunjukkan pada sumbu mendatar atau sumbu kemampuan (θ) tersebut merupakan nilai parameter b (tingkat kesukaran) dari TCC. Berdasarkan kurva karakteristik total, diketahui bahwa nilai parameter b (tingkat kesukaran) dari soal penilaian akhir semester fisika kelas X adalah sebesar $-1,4$ atau berada dalam kategori tingkat kesukaran mudah.

Nilai parameter a (daya pembeda) dapat diperoleh dari nilai kemiringan (*slope*) atau diperoleh dari nilai $\tan \alpha$. Berdasarkan kurva karakteristik total, diperoleh bahwa soal penilaian akhir semester fisika kelas X memiliki nilai parameter a (daya pembeda) sebesar 0,354 dan dikategorikan baik karena nilai parameter a berada pada rentang 0 sampai dengan 2.

Parameter Butir Soal

Estimasi parameter butir soal yang terdiri dari daya pembeda (a), tingkat kesukaran (b), dan *Standard Error* atau kesalahan pengukuran ($s.e$) yang diperoleh dari hasil analisis soal penilaian akhir semester Fisika kelas X berdasarkan model 2 parameter logistik disajikan pada tabel 2 berikut.

Tabel 2 Hasil estimasi parameter butir soal model 2 parameter logistik (2 PL)

Butir	Daya Pembeda (a)	s.e	Kriteria	Tingkat Kesukaran (b)	s.e	Kriteria
1	2,155	0,443	Tidak baik	-2,091	0,214	Tidak baik
2	1,122	0,341	Baik	-2,751	0,572	Tidak baik
3	0,197	0,206	Baik	-2,660	2,224	Tidak Baik
4	1,406	0,255	Baik	-0,687	0,107	Baik
5	1,702	0,000	Baik	0,000	0,000	Baik
6	2,045	0,372	Tidak baik	-1,771	0,162	Baik
7	2,598	0,477	Tidak baik	-1,895	0,153	Baik
8	0,309	0,282	Baik	5,044	5,122	Tidak baik
9	1,397	0,275	Baik	-1,499	0,174	Baik
10	3,255	0,568	Tidak baik	-1,831	0,121	Baik
11	2,140	0,356	Tidak baik	-1,538	0,126	Baik
12	1,702	0,000	Baik	0,000	0,000	Baik
13	3,677	0,553	Tidak baik	-1,523	0,082	Baik
14	3,484	0,582	Tidak baik	-1,750	0,106	Baik
15	0,123	0,244	Baik	9,878	20,710	Tidak baik

16	2,489	0,392	Tidak baik	-1,486	0,107	Baik
17	1,993	0,378	Baik	-1,862	0,180	Baik
18	2,204	0,441	Tidak baik	-2,040	0,200	Tidak baik
19	1,225	0,244	Baik	-0,589	0,120	Baik
20	1,906	0,308	Baik	-1,254	0,110	Baik
21	1,156	0,252	Baik	-1,436	0,196	Baik
22	2,386	0,428	Tidak baik	-1,815	0,150	Baik
23	3,802	0,581	Tidak baik	-1,569	0,084	Baik
24	2,279	0,437	Tidak baik	-1,953	0,179	Baik
25	0,241	0,296	Baik	7,256	9,591	Tidak baik

Berdasarkan tabel 2, terdapat 13 butir soal yang memiliki estimasi daya pembeda (a) pada rentang 0 sampai 2, sedangkan 12 butir soal lainnya memiliki estimasi daya pembeda (a) di atas 2 yaitu butir soal ke-1, ke-6, ke-7, ke-10, ke-11, ke-13, ke-14, ke-16, ke-18, ke-22, ke-23, dan ke-24. Hal ini dapat ditafsirkan bahwa terdapat 13 butir soal penilaian akhir semester fisika kelas X yang memiliki estimasi daya pembeda dengan kategori baik dan terdapat 12 butir soal penilaian akhir semester fisika kelas X yang memiliki estimasi daya pembeda dengan kategori tidak baik. Sementara itu, untuk estimasi tingkat kesukaran terdapat 18 butir soal penilaian akhir semester fisika kelas X yang dikategorikan baik dikarenakan memiliki estimasi tingkat kesukaran (b) pada rentang -2 sampai 2, sedangkan 8 butir soal yakni butir ke-1, ke-2, ke-3, ke-8, ke-15, ke-18, dan ke-25 dikategorikan tidak baik dikarenakan memiliki estimasi tingkat kesukaran (b) di bawah -2 dan di atas 2.

Kesimpulan

Dari hasil penelitian yang dilakukan dapat disimpulkan bahwa terdapat 2 butir soal penilaian akhir semester fisika kelas X yang tidak valid (butir ke-5 dan ke -12) dan terdapat 23 butir soal yang valid. Selain itu, soal penilaian akhir semester fisika kelas X reliabel jika diujikan pada peserta didik dengan rentang kemampuan -3,619 sampai 1,206 yaitu kemampuan sangat rendah hingga tinggi. Secara keseluruhan, soal penilaian akhir semester fisika kelas X memiliki nilai parameter a (daya pembeda) sebesar 0,354 yang dapat dikategorikan baik karena nilai parameter a berada pada rentang 0 sampai dengan 2 dan nilai parameter b (tingkat kesukaran) dari soal penilaian akhir semester fisika kelas X adalah sebesar -1,4 atau berada dalam kategori tingkat kesukaran mudah.

Daftar Pustaka

- [1] H. L. Astuti, "Analisis Soal Tes Penilaian Akhir Semester Fisika Kelas X Menggunakan Teori Respon Butir," S.Pd. Skripsi, Program Studi Pendidikan Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Yogyakarta, 2019.
- [2] J. M. Azevedo, E.P Oliveira, dan P.D Beites, "Using Learning Analytics to Evaluate The Quality of Multiple-Choice Questions A Perspective With Classical Test Theory and Item Response Theory," *International Journal of Information and Learning Technology*, vol. 36, no. 4, hlm. 322-341, Feb 2019, doi: 10.1108/IJILT-02-2019-0023.
- [3] P. Eaton, dkk., "Classical Test Theory and Item Response Theory Comparison of The Brief Electricity and Magnetism Assessment and The Conceptual Survey of Electricity and Magnetism," *Physical Review Physics Education Research*, vol. 15, no. 1, hlm. 1-19, 2019. doi: 10.1103/PhysRevPhysEduRes.15.010102.
- [4] R.K Hambleton, H. Swaminathan, dan H.J Rogers, *Fundamentals of Item Response Theory*, Newbury Park, California: Sage Publications Inc, 1991.
- [5] Kemendikbud, *Permendikbud No. 23 tentang Standar Penilaian Pendidikan*, Jakarta: Kementerian Pendidikan dan Kebudayaan, 2016.

- [6] E. Ratnawulan dan Rusdiana, *Evaluasi Pembelajaran dengan Pendekatan Kurikulum 2013*, Bandung: Pustaka Setia, 2014.
- [7] H. Retnawati, *Teori Respons Butir dan Penerapannya*, Yogyakarta: Nuha Medika, 2014.
- [8] A. Sopiah, S. Sidauruk, N.B Asi, " Kualitas Soal Penilaian Akhir Semester (PAS) Buatan Guru Mata Pelajaran Kimia Kelas X IPA SMA Negeri Di Kabupaten Seruyan Pada Semester Ganjil tahun Ajaran 2018/2019", *Jurnal Ilmiah Kanderang Tingang*, vol. 10, no. 2, hlm. 110-126, 2019.