

# VALIDASI PERSEPSI SISWA TERHADAP PEMBELAJARAN SAINS DENGAN TEORI RESPON BUTIR MULTIDIMENSI

Janu Arlinwibowo<sup>a,\*</sup>, Samsul Hadi<sup>b</sup>, Eric Maurteen Firdaus<sup>a</sup>

<sup>a</sup>[janu@umkudus.ac.id](mailto:janu@umkudus.ac.id), <sup>b</sup>[samsulhd@gmail.com](mailto:samsulhd@gmail.com), <sup>c</sup>[erikmaurteen@umkudus.ac.id](mailto:erikmaurteen@umkudus.ac.id)

<sup>a,c</sup>Fakultas Sains, Teknologi, dan Matematika, Universitas Muhammadiyah Kudus

Jl. Ganesha Raya No.I, Purwosari, Kabupaten Kudus, Jawa Tengah 59316, Indonesia

<sup>b</sup>Fakultas Teknik, Universitas Negeri Yogyakarta

Jl. Colombo No.1, Karang Malang, Caturtunggal, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta 55281, Indonesia

## Abstrak

Penelitian itu bertujuan untuk melakukan re-validasi angket persepsi siswa terhadap pembelajaran IPA dalam survei PISA 2015, yang diduga multidimensi, dengan MIRT untuk melihat dimensi tes dan parameter butir. Proses analisis diharapkan dapat menghasilkan suatu kesimpulan kualitas item dan tes secara lebih presisi. Penelitian ini akan mengkaji pola respon 5870 siswa Indonesia yang mengisi 23 butir pertanyaan pilihan ganda bergradasi mulai dari 1 hingga 4. Proses penelitian dimulai dengan analisis berbasis Principal Component Analysis (rotasi varimax), uji fit model dengan signed chi-squared test, melakukan estimasi parameter berdasar pada pembagian faktor sesuai dengan hasil PCA, dan penyimpulan hasil analisis. Angket persepsi siswa terhadap pembelajaran IPA terdiri dari 23 item yang memuat multidimensi (5 faktor). Diantaranya terdapat beberapa butir yang overlapping (butir yang mengukur lebih dari satu faktor) yaitu butir 4, 8, 10, 14, 19, 20. Berdasar pada MDISC, semua item berada pada level minimal yaitu moderate dengan arti bahwa item mampu dengan baik membedakan siswa berdasarkan kemampuannya. Berdasarkan MDIFF, semua item memiliki parameter step yang runtut mulai dari kecil hingga besar sehingga MDIFF mampu mendefinisikan kemampuan siswa secara baik.

**Kata Kunci:** Validasi, Persepsi Siswa, Pembelajaran Sains, Teori Respon Butir, Multidimensi, PISA 2015

## Abstract

*The research aims to re-validate the students' perception questionnaire towards natural science learning in the 2015 PISA survey, which is suspected to be multidimensional, with MIRT to see the dimensions of the test and item parameters. The analysis process is expected to produce a conclusion about the quality of items and tests more precisely. This study will examine the response patterns of 5870 Indonesian students who fill 23 graded multiple-choice questions ranging from 1 to 4. The research process begins with Principal Component Analysis (varimax rotation) analysis, model fit test with the signed chi-squared test, parameter estimation based on the division of factors in accordance with the results of the PCA, and the conclusion of the analysis. Questionnaire students' perceptions of science learning consist of 23 items that contain multidimensional (5 factors). Among them, there are several overlapping items (items that measure more than one factor) namely items 4, 8, 10, 14, 19, 20. Based on MDISC, all items are at a minimal level that is moderate with the meaning that items are able to properly distinguish students based on his abilities. Based on MDIFF, all items have coherent step parameters ranging from small to large so that MDIFF is able to define students' abilities well.*

**Keywords:** Validation, Student Perception, Learning Science, Item Response Theory, Multidimensional, PISA 2015

## I. PENDAHULUAN

Survei internasional PISA (Program for International Student Assessment), like other international studies such as TIMSS (Trends in Mathematics and Science Study) and PIRLS (Progress in International Reading

Literacy Study), telah dilakukan oleh peneliti-peneliti di Dunia. Tujuan dari survei-survei tersebut adalah untuk mengukur kemampuan dan ketrampilan siswa terutama terkait dengan masalah keseharian (Ceylan and Serdar Abaci, 2013). Dengan demikian maka akan ada suatu peta dunia terkait pendidikan

secara jelas yang kemudian dapat dijadikan sarana untuk refleksi diri.

Survei PISA 2015 berfokus pada sains, dengan membaca, matematika dan pemecahan masalah kolaboratif sebagai bidang penilaian minor. PISA 2015 juga menyertakan penilaian literasi keuangan kaum muda, yang merupakan pilihan untuk negara dan ekonomi (OECD, 2018). Salah satu data hasil survei dari PISA adalah persepsi siswa terhadap pembelajaran IPA. Data persepsi ini menjadi menarik karena menunjukkan bahwa ada suatu hubungan positif antara persepsi dengan prestasi belajar (Mayya & Roff, 2004; Ahmed, Taha, Alneel, & Gaffar, 2018). Persepsi siswa terhadap suatu pelajaran dapat dijadikan suatu gambaran situasi pendidikan dan juga sebagai tolak ukur kualitasnya.

Salah satu syarat kesuksesan suatu penelitian adalah kualitas instrumen. Jika instrumen baik, maka kemungkinan besar suatu data hasil penelitian dapat memberikan gambaran yang relatif riil dan objektif. Dengan demikian berdasar fakta urgensi data persepsi siswa dalam memotrek keadaan pendidikan suatu negara maka menjadi hal penting untuk dilakukan suatu kajian terhadap validasi instrumen survei persepsi siswa terhadap pelajaran IPA untuk memastikan bahwa data yang terhimpun merepresentasikan kondisi nyata.

Pada umumnya analisis yang dilakukan adalah berbasis pada item response theory unidimensi sehingga menurut Hambleton & Swaminathan (1985) suatu tes harus hanya mengukur satu kemampuan. Namun demikian ada suatu syarat yang relatif sulit dipenuhi yaitu bahwa data empiris harus menunjukkan unidimensi. Namun faktanya, suatu tes sangat kompleks dan cenderung mengukur berbagai hal sehingga banyak kejadian data empirik menunjukkan bahwa suatu tes mengelompok dalam beberapa faktor (multidimensi).

Embretson & Reise (2000) menyatakan bahwa untuk melakukan analisis tes yang terbukti mengukur lebih dari satu kemampuan menggunakan multidimensi item response theory (MIRT). Rackase (2009) menyampaikan bahwa model logistik MIRT adalah sebagai berikut:

$$P(X = 1 | \theta, a, d) = \frac{e^{a \cdot \theta^T + d}}{1 + e^{a \cdot \theta^T + d}},$$

Dimana  $\theta = \theta_1, \theta_2, \theta_3, \dots, \theta_k$  adalah vektor dari kemampuan siswa dan  $a = a_1, a_2, a_3, \dots, a_k$  adalah vektor dari diskriminasi item serta  $d$  adalah tingkat kesulitan item. Kemampuan diskriminasi dari item untuk kombinasi berbagai diskriminasi dapat disimpulkan dengan rumus sebagai berikut:

$$MDISC = \sqrt{\sum_{j=1}^k a_j^2}$$

Tingkat kesulitan dari setiap item tes dapat dikalkulasi dengan rumus sebagai berikut:

$$MDIFF = -\frac{d}{MDISC}$$

Dengan demikian maka akan dilakukan revalidasi angket persepsi siswa terhadap pelajaran IPA dalam survei PISA 2015, yang diduga multidimensi, dengan MIRT untuk melihat dimensi tes dan parameter butir. Proses analisis diharapkan dapat menghasilkan suatu kesimpulan kualitas item dan tes secara lebih presisi.

## II. METODE PENELITIAN

### A. Definisi Data

Sekitar 540.000 siswa terlibat dalam survei PISA di 2015, mewakili sekitar 29 juta anak berusia 15 tahun di sekolah-sekolah dari 72 negara yang berpartisipasi. Indonesia merupakan salah satu negara yang mengikuti survei PISA pada tahun 2015. Penelitian ini melakukan revalidasi terhadap instrumen PISA yang digunakan untuk mengukur persepsi siswa terhadap pelajaran IPA. Peneliti memilih data di Indonesia yang melibatkan 6513 siswa. Namun demikian dilakukan pemilihan data yang lengkap sehingga yang masuk dalam analisis adalah 5870 siswa Indonesia.

### B. Instrumen

Instrumen yang digunakan dalam penelitian ini adalah instrumen siswa berupa angket yang menilai mengenai pembelajaran IPA. Diambil 23 item terkait persepsi siswa terhadap pembelajaran IPA disekolahnya (ST098, ST100, ST103, dan ST104). Instrumen merupakan pilihan ganda dengan 4

pilihan jawaban yang menunjukkan gradasi yang diberi poin mulai 1 sampai dengan 4. Dengan demikian maka analisis multidimensi akan dilakukan dengan menggunakan graded response model (GRM)

**Analisis**

Proses pertama yang dilakukan adalah analisis berbasis Principal Component Analysis (PCA) dengan rotasi varimax untuk menunjukkan pembagian faktor dari data empiri yang terhimpun. Langkah kedua adalah melakukan uji fit model dengan memanfaatkan software R packages mirt berbasis pada signed chi-squared test. Dari hasil uji fit model tersebut akan dihasilkan p-value dan RMSEA dengan ketentuan data fit model jika  $RMSEA < 0.08$ . Langkah ketiga adalah melakukan estimasi parameter berdasar pada pembagian faktor sesuai dengan hasil PCA dengan model GRM hingga menghasilkan MDIFF dan MDISC. Langkah terakhir adalah penyimpulan hasil analisis.

**III. HASIL DAN PEMBAHASAN**

**C. Penentuan Dimensi Berbasis PCA**

Dalam melakukan analisis data dengan menggunakan teori respon butir, hal yang pertama yang dilakukan adalah menguji dimensi dari data empiris yang di dapatkan. Proses pengujian dilakukan dengan analisis faktor eksploratori dengan metode principal componen. Asumsi awal yang harus dipenuhi dalam analisis faktor eksploratori adalah tes KMO dan Bartlett sebagai berikut.

**Tabel 1.** KMO and Bartlett's Test

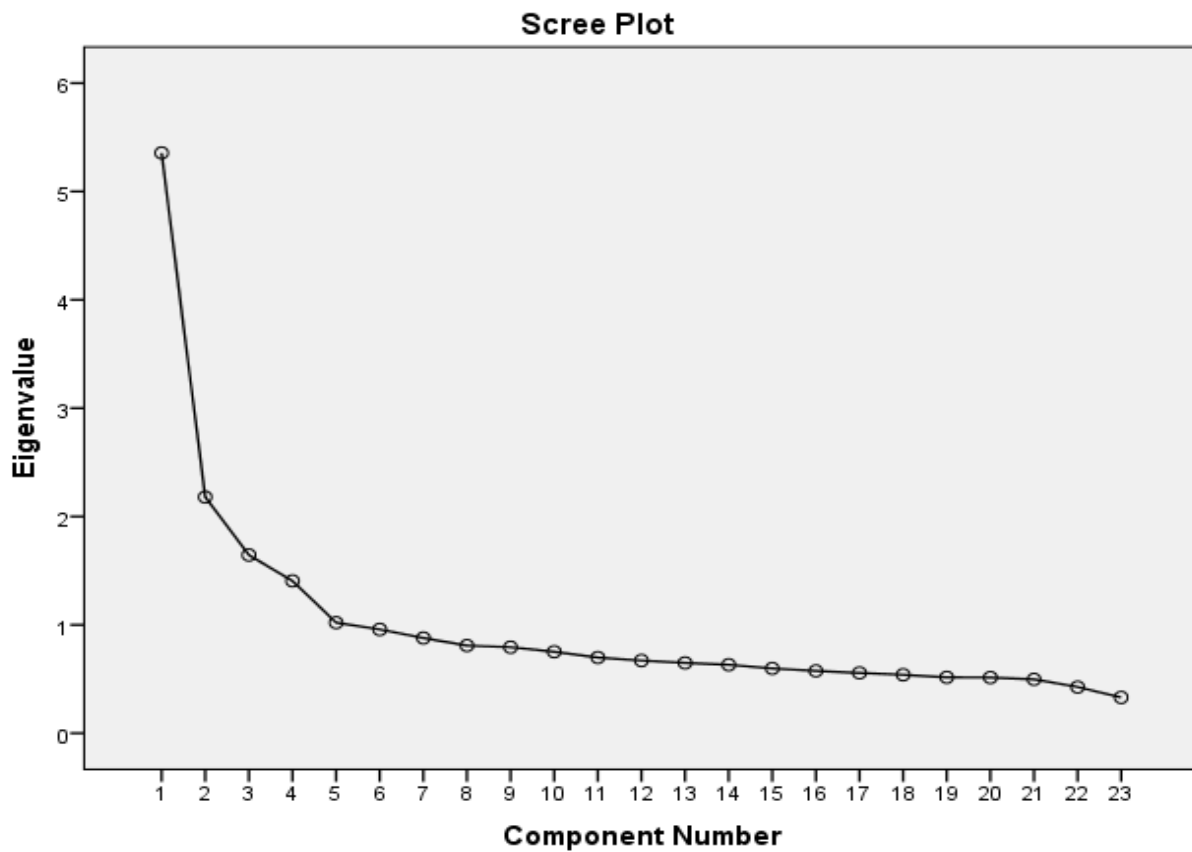
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.875
Bartlett's Test of Sphericity	Approx. Chi-Square	32191.489
	df	253
	Sig.	0.000

Dalam penelitian ini, dapat dikatakan telah memiliki kecukupan data yang diindikasikan dengan  $KMO > 0,5$ . Selain itu signifikansi Bartlett's tes menunjukkan bahwa Hipotesis  $H_0$  matriks korelasi merupakan matriks identitas ditolak sehingga data membentuk matriks korelasi dengan ada hubungan kedekatan antar variabel. Kemudian bersakan analisis pada semua data anti-images correlation maka measures of adequate sampling (MSA) telah memiliki nilai diatas 0,5 sehingga semua data telah layak masuk dalam analisis faktor. Dalam analisis faktor dihasilkan data pengelompokan sebagai berikut.

**Tabel 2.** Eigen Value hasil PCA

C	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	5.356	23.287	23.287
2	2.179	9.476	32.762
3	1.645	7.150	39.913
4	1.406	6.112	46.025
5	1.021	4.438	50.463
...	...	...	...
...	...	...	...
23	.330	1.434	100.000

Analisis faktor dengan metode Principal Component Analysis (PCA) mereduksi atau meminimalkan jumlah variabel amatan agar menjadi komponen-komponen utama yang jumlahnya lebih sedikit dibandingkan dengan jumlah butir. Komponen utama yang kemudian menjadi faktor ditentukan dengan memilih komponen-komponen yang memiliki nilai eigen lebih dari 1. Nilai tersebut adalah acuan untuk menghitung faktor pengelompokan yang diketahui serta pengaruhnya ditunjukkan dari komulatif varian (Guttman, 1954; Kaiser, 1960). Berikut adalah paparan scree plot yang mendefinisikan nilai eigen.



Gambar 1. Scree Plot

Dengan demikian data empiris menunjukkan bahwa item butir persepsi siswa terhadap pembelajaran ipa pada survei PISA terbagi menjadi 5 faktor. Dengan

menggunakan rotasi varimax maka 23 butir pertanyaan mengelompok dalam 5 faktor sebagai berikut.

Tabel 3. Hasil Principal Component Analysis

Item	Component				
	1	2	3	4	5
Siswa diberi kesempatan untuk menjelaskan pendapatnya				.608	
Siswa melakukan praktek percobaan di laboratorium		.576			
Siswa diharapkan untuk mengemukakan alasan mengenai pertanyaan IPA.				.682	
Siswa diminta untuk membuat kesimpulan dari percobaan yang telah mereka lakukan.		.462		.516	
Guru menjelaskan cara suatu ide dalam IPA dapat diterapkan pada sejumlah fenomena yang berbeda (misal, gerakan benda, bahan-bahan dengan sifat-sifat yang serupa).				.531	
Siswa diperbolehkan merencanakan percobaannya sendiri.		.611			
Terdapat kelas debat mengenai investigasi.		.677			
Guru menjelaskan dengan baik hubungan antara konsep-konsep IPA dengan kehidupan kita.			.377	.474	
Siswa diminta untuk melakukan investigasi untuk menguji idenya.		.583			
Guru tertarik pada kegiatan belajar setiap siswa.		.404	.473		
Guru memberi bantuan tambahan bagi siswa yang membutuhkannya.			.686		

Guru membantu siswa dalam belajar.	.717		
Guru terus mengajar sampai siswa mengerti.	.684		
Guru memberi kesempatan kepada siswa untuk mengemukakan pendapatnya.	.561	.356	
Guru menjelaskan ide-ide ilmiah.			.608
Terjadi diskusi seluruh siswa di kelas dengan guru.			.730
Guru mendiskusikan pertanyaan-pertanyaan kami.			.746
Guru mendemonstrasikan sebuah ide.			.697
Guru memberi tahu saya mengenai prestasi saya di pelajaran IPA.	.602	-.315	
Guru saya memberi masukan mengenai kelebihan saya dalam mata pelajaran IPA.	.665	-.329	
Guru memberi tahu saya materi apa saja yang masih bisa saya tingkatkan.	.744		
Guru memberitahu saya bagaimana cara meningkatkan prestasi saya.	.775		
Guru memberi saran kepada saya bagaimana cara meraih tujuan pembelajaran saya.	.748		

Berbasis pada analisis faktor maka dapat disimpulkan bahwa data tidak unidimensi sehingga tidak dapat dilakukan analisis teori respon butir unidimensi. Oleh karena itu data kemudian akan dianalisis dengan menggunakan teori respon butir multidimensi, dalam kasus ini memuat 5 dimensi. Berikut adalah rangkuman dimensi (faktor) dan anggota himpunannya (butir).

- F1 = butir ke- 19, 20, 21, 22, 23
- F2 = butir ke- 2, 4, 5, 6, 8, 9, 19, 20
- F3 = butir ke- 8, 10, 11, 12, 13, 14
- F4 = butir ke- 1, 3, 4, 5, 8, 14
- F5 = butir ke- 15, 16, 17, 18

Nampak bahwa terdapat beberapa butir yang overlapping (butir yang mengukur lebih dari satu faktor) yaitu butir 4, 8, 10, 14, 19, 20.

### Analisis Respon Butir Multidimensi

Tabel 4. Estimasi Parameter Butir GRM

	a1	a2	a3	a4	a5	d1	d2	d3
B1	0.000	0.000	0.000	1.744	0.000	3.967	0.542	-0.896
B2	0.000	0.900	0.000	0.000	0.000	0.821	-2.591	-3.739
B3	0.000	0.000	0.000	2.124	0.000	3.381	-0.004	-1.811
B4	0.000	0.788	0.000	0.975	0.000	2.837	-0.595	-2.161
B5	0.000	0.000	0.000	1.278	0.000	3.189	0.358	-0.957
B6	0.000	1.404	0.000	0.000	0.000	1.122	-1.731	-3.032
B7	0.000	1.161	0.000	0.000	0.000	-0.005	-2.239	-3.539
B8	0.000	0.000	0.829	0.925	0.000	3.845	1.155	-0.468
B9	0.000	1.477	0.000	0.000	0.000	1.643	-1.304	-2.766
B10	0.000	0.539	0.736	0.000	0.000	2.831	-0.196	-1.566
B11	0.000	0.000	1.602	0.000	0.000	3.842	1.004	-0.712

Analisis diawali dengan menguji kecocokan data dengan model estimasi parameter butir. Uji kecocokan memanfaatkan package “mirt” yang menghasilkan luaran RMSEA dan p-value untuk mengindikasikan suatu kecocokan. Kecocokan model dari data ujian nasional dilakukan dengan memperhatikan nilai RMSEA saja karena p-value sensitif terhadap jumlah sampel (pada umumnya semakin besar sampel maka nilai cenderung mengecil sehingga  $H_0$  ditolak). Dengan ketentuan bahwa model cocok jika  $RMSEA < 0.08$  maka model estimasi parameter butir memiliki kecocokan dengan model estimasi parameter graded response model (GRM). Oleh karena itu, selanjutnya analisis yang digunakan akan berbasis pada model GRM. Hasil dari estimasi parameter soal dengan multidimensi teori respon butir adalah sebagai berikut.

B12	0.000	0.000	1.783	0.000	0.000	4.412	1.924	0.212
B13	0.000	0.000	1.780	0.000	0.000	5.054	2.486	0.648
B14	0.000	0.000	1.398	0.997	0.000	5.227	1.902	-0.017
B15	0.000	0.000	0.000	0.000	1.262	1.941	0.689	-2.831
B16	0.000	0.000	0.000	0.000	1.751	2.322	0.709	-3.470
B17	0.000	0.000	0.000	0.000	1.952	2.372	0.616	-3.070
B18	0.000	0.000	0.000	0.000	1.648	2.481	0.945	-2.540
B19	1.034	-0.396	0.000	0.000	0.000	3.084	1.809	-1.064
B20	1.243	-0.456	0.000	0.000	0.000	3.291	1.880	-0.958
B21	1.952	0.000	0.000	0.000	0.000	3.075	1.182	-2.107
B22	2.763	0.000	0.000	0.000	0.000	2.671	0.234	-4.069
B23	2.402	0.000	0.000	0.000	0.000	2.091	0.030	-4.011

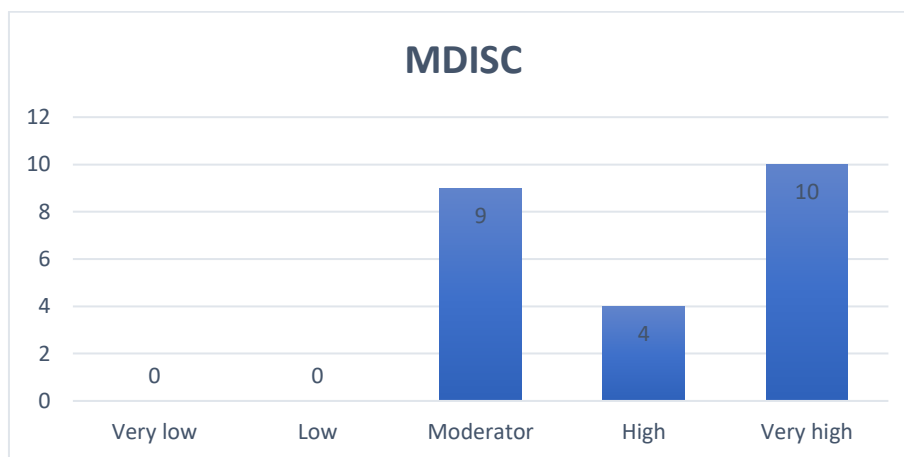
Diskriminan atau sering juga disebut slope ditandai oleh suatu kemiringannya. Kemiringan positif menunjukkan bahwa kemungkinan respons yang benar dari siswa yang baik lebih tinggi daripada siswa yang buruk, sedangkan kemiringan negatif menggambarkan tren yang berlawanan. Untuk analisis lebih lanjut, kombinasi diskriminatif dan kesulitan soal yang disebutkan dalam rumus MDISC dan MDIFF harus dihitung.

MDISC adalah singkatan dari kombinasi diskriminatif, dan MDIFF menunjukkan kesulitan item. Menurut Baker (2001) dan Hasmy (2014), kombinasi yang membedakan dan kesulitan item dapat diklasifikasikan masing-masing sebagai berikut:

Tabel 5. Kasifikasi MDISC

Kriteria	Keterangan
$MDISC \geq 1.7$	Very high
$1.35 \leq MDISC < 1.7$	High
$0.65 \leq MDISC < 1.35$	Moderator
$0.35 \leq MDISC < 0.65$	Low
$MDISC < 0.35$	Very low

Semua item memiliki diskriminasi yang cukup baik yaitu minimal memiliki indeks diskriminasi kategori moderate yaitu 9 masuk kategori moderate, 4 masuk kategori high, dan 10 masuk kategori very high. Jika berbasis pada klasifikasi MDISC maka semua butir telah masuk kategori baik sehingga tidak ada yang harus dikeluarkan. Berikut adalah sebaran butir menurut MDISC.



Gambar 1. Sebaran MDISC

Mengenai kesulitan item, analisis politomus menghasilkan tingkat kesulitan yang dapat diartikan pula dengan parameter step dimana tingkat kesulitan tersebut dapat mendefinisikan kemampuan minimal seorang siswa untuk dapat memilih suatu jawaban tertentu. Hal yang menjadi perhatian adalah

tingkat kesulitan adalah bahwa tingkat kesulitan dapat memberikan definisi kemampuan siswa dengan baik saat tidak ada kecacauan posisi tingkat kesulitan 1 dengan tingkat kesulitan lain. Formasi yang harusnya muncul pada suatu item soal dengan 4 pilihan jawaban adalah  $d1 < d2 < d3$ . Artinya

adalah dengan kemampuan kurang dari  $d_1$  maka siswa berpeluang besar menjawab pilihan a, siswa kemungkinan akan menjawab pilihan b dengan memiliki minimal kemampuan  $d_1$ , siswa akan menjawab c dengan memiliki minimal kemampuan  $d_2$ , dan siswa akan menjawab d dengan memiliki minimal kemampuan  $d_3$ . Hal tersebut tentunya berlaku untuk pilihan jawaban yang berjenjang yaitu  $a < b < c < d$ .

Pada faktanya banyak data yang menunjukkan inkonsistensi jawaban sehingga aturan tersebut tidak dapat terpenuhi. Posisi tingkat kesulitan tidak terurut dengan baik sehingga definisi kemampuan siswa sulit (tidak bisa) untuk diamati. Dengan demikian maka akan ditelusuri kesesuaian hasil perhitungan MDIFF dengan aturan tersebut.

**Tabel 6.** MDIFF

Item	d1	Kesesuaian	d2	Kesesuaian	d3
B1	-2.274	TRUE	-0.311	TRUE	0.514
B2	-0.913	TRUE	2.880	TRUE	4.155
B3	-1.592	TRUE	0.002	TRUE	0.853
B4	-2.262	TRUE	0.474	TRUE	1.724
B5	-2.495	TRUE	-0.280	TRUE	0.749
B6	-0.799	TRUE	1.233	TRUE	2.160
B7	0.005	TRUE	1.929	TRUE	3.048
B8	-3.096	TRUE	-0.930	TRUE	0.377
B9	-1.112	TRUE	0.883	TRUE	1.873
B10	-3.102	TRUE	0.215	TRUE	1.716
B11	-2.398	TRUE	-0.627	TRUE	0.445
B12	-2.474	TRUE	-1.079	TRUE	-0.119
B13	-2.839	TRUE	-1.397	TRUE	-0.364
B14	-3.044	TRUE	-1.108	TRUE	0.010
B15	-1.537	TRUE	-0.546	TRUE	2.242
B16	-1.326	TRUE	-0.405	TRUE	1.981
B17	-1.215	TRUE	-0.315	TRUE	1.573
B18	-1.506	TRUE	-0.573	TRUE	1.541
B19	-2.784	TRUE	-1.633	TRUE	0.960
B20	-2.485	TRUE	-1.420	TRUE	0.724
B21	-1.575	TRUE	-0.605	TRUE	1.079
B22	-0.967	TRUE	-0.085	TRUE	1.472
B23	-0.871	TRUE	-0.013	TRUE	1.670

Berdasarkan data perhitungan, semua butir mampu mendefinisikan kemampuan siswa dengan baik. Hal tersebut ditunjukkan dengan label “TRUE” pada kolom kesesuaian yang menandakan bahwa  $d_n > d_{n-1}$  sehingga perpotongan antar kurva karakteristik tersusun dengan baik sehingga parameter step dapat berfungsi mendefinisikan kemampuan siswa secara baik. Berdasarkan temuan tersebut maka dapat disimpulkan bahwa menurut MDIFF tidak ada butir yang harus dikeluarkan.

**IV. KESIMPULAN**

Angket persepsi siswa terhadap pembelajaran IPA terdiri dari 23 item yang memuat multidimensi (5 faktor). Diantaranya

terdapat beberapa butir yang overlapping (butir yang mengukur lebih dari satu faktor) yaitu butir 4, 8, 10, 14, 19, 20. Berdasar pada MDISC, semua item berada pada level minimal yaitu moderate dengan arti bahwa item mampu dengan baik membedakan siswa berdasarkan kemampuannya. Berdasarkan MDIFF, semua item memiliki parameter *step* yang runtut mulai dari kecil hingga besar sehingga MDIFF mampu mendefinisikan kemampuan siswa secara baik.

**DAFTAR PUSTAKA**

Ahmed, Y., Taha, M. H., Alneel, S., & Gaffar, A. M. (2018). Students’ perception of the learning environment and its relation to their study year and performance in

- Sudan. *International Journal of Medical Education*, 2018(9), 145-150.
- Baker, F. (2001). *The basic of item response theory*. USA: ERIC Clearinghouse on Assessment and Evaluation.
- Ceylan, E & Abaci, S. (2013). Differences between Turkey and Finland based on Eight Latent Variables in PISA 2006. *International Online Journal of Educational Sciences*, 5 (1), 10-21.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. USA: Kluwer-Nijhoff Publishing.
- Hasmy, A. (2014). Compare unidimensional & multidimensional Rasch model for test with multidimensional construct and items local dependence. *Journal of Education and Learning*, 8(3), 187-194.
- Mayya, S. S. & Roff, S. (2004). Students' perceptions of educational environment: a comparison of academic achievers and under-achievers at Kasturba Medical College, India. *Educ Health*. 2004 (17), 280-291.
- OECD. (2018). *PISA 2015 Results*. PISA in Focus, OECD Publishing.
- Reckase, M. D. (2009). *Multidimensional Item Response Theory*. NY: Springer