



Karakterisasi instrumen tes keterampilan berpikir kritis dengan analisis model *rasch* pada materi alat optik

Gini Erdiani*, W. Liliawati, Muslim

Program Studi Pendidikan Fisika, Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam, Universitas Pendidikan Indonesia

*e-mail: ginierdiani97@gmail.com

Abstrak

Keterampilan berpikir kritis (KBK) merupakan salah satu keterampilan yang harus dimiliki siswa di abad ke-21, karena telah menjadi salah satu kompetensi dalam tujuan pendidikan. Untuk mengetahui keterampilan berpikir kritis siswa tentunya diperlukan suatu instrumen tes dengan karakteristik yang sesuai. Penelitian ini dilakukan dengan tujuan untuk mengkarakterisasi tes keterampilan berpikir kritis berdasarkan analisis model Rasch dengan bantuan aplikasi Ministep. Desain penelitian yang digunakan adalah desain konstruksi dan validasi tes. Proses konstruksi menghasilkan tes KBK berbentuk pilihan ganda yang mengukur 5 aspek KBK. Sedangkan, proses validasi didasarkan pada penilaian 5 orang ahli dan uji coba tes oleh 75 partisipan dari SMA yang ada di Bandung. Hasil penelitian menunjukkan bahwa validitas isi tes KBK berdasarkan analisis CVR (Content Validity Ratio) dikategorikan valid. Selain itu hasil validitas empiris tes KBK menunjukkan bahwa terdapat 23 soal yang valid. Reliabilitas tes memiliki reliabilitas dalam kategori cukup. Tingkat kesukaran tes KBK menunjukkan soal dengan kriteria sukar terdapat pada soal S6 dengan nilai logit +1,31. Sedangkan, soal yang termasuk kategori mudah terdapat pada soal S12 dengan nilai logit -2,74. Pada instrumen tes KBK yang dikonstruksi masih terdapat beberapa soal yang bias terhadap jenis kelamin tertentu yaitu soal S2, S8, S10, S16, S18, S19, S22, dan S23.

Kata kunci: model Rasch; karakterisasi tes; keterampilan berpikir kritis

1. Pendahuluan

Berpikir kritis telah menjadi suatu istilah yang 'sangat populer' dalam dunia pendidikan di abad ke-21 (Fisher, 2009). Berpikir kritis dapat diartikan sebagai proses dan kemampuan yang digunakan untuk memahami konsep, menerapkan, mensintesis dan mengevaluasi informasi yang diperoleh atau informasi yang dihasilkan (Zubaidah, 2010). Di dunia saat ini, orang mengakses sejumlah besar informasi dengan cepat, namun pada dasarnya tidak semua informasi yang didapatkan bisa diyakini dan dijadikan landasan dalam bertindak karena tidak selalu informasi yang didapatkan tersebut merupakan informasi yang benar. Oleh karena itu mereka membutuhkan keterampilan berpikir lanjut untuk mengkategorikan informasi sebagai benar dan tidak benar untuk pekerjaan atau pengambilan keputusan (Oonsim dan Chanprasert, 2017).

Keterampilan berpikir kritis perlu ditanamkan pada diri setiap siswa untuk menanggapi tantangan hidup yang sangat

kompleks (Halpern, 2003 dalam Wartono dkk, 2018). Sejalan dengan pendapat Wagner (2010), ia mengidentifikasi ada sedikitnya tujuh kompetensi dan keterampilan bertahan hidup yang perlu dimiliki oleh siswa dalam menghadapi kehidupan dan dunia kerja di abad ke-21, yaitu: (1) kemampuan berpikir kritis dan pemecahan masalah, (2) kolaborasi dan kepemimpinan, (3) ketangkasan dan kemampuan beradaptasi, (4) inisiatif dan berjiwa entrepreneur, (5) mampu berkomunikasi efektif baik secara oral maupun tertulis, (6) mampu mengakses dan menganalisis informasi, dan (7) memiliki rasa ingin tahu dan imajinasi.

Keterampilan berpikir kritis memainkan peranan yang penting dalam semua aspek kehidupan manusia terutama aspek pendidikan. Oleh karena itu, keterampilan berpikir kritis diperlukan untuk terus dilatihkan dalam proses pembelajaran (Kealey dkk, 2005 dalam Wartono dkk 2018). Dengan tujuan agar menghasilkan peserta didik yang kompeten dan terampil dalam

memecahkan masalah di kehidupan sehari-hari. Keterampilan berpikir kritis tidak hanya diterapkan dalam pembelajaran saja tetapi juga harus di dukung dengan instrumen penilaian yang dapat mengukur tingkat keterampilan berpikir kritis yang dimiliki siswa.

Instrumen tes fisika di sekolah hanya mengukur keterampilan berpikir dasar siswa (Sugiarti dkk, 2017). Belum banyak digunakan untuk mengukur keterampilan berpikir kritis yang termasuk salah satu keterampilan berpikir tingkat tinggi. Padahal, alat penilaian yang digunakan oleh guru memainkan peran penting dalam menilai prestasi siswa dan pengembangan keterampilan berpikir kritis siswa (Ozsevgec, 2006).

Dengan adanya alat penilaian berupa instrumen tes dengan karakteristik tertentu yang mampu mengukur keterampilan berpikir kritis siswa, diharapkan dapat memberikan informasi yang akurat yang dapat menggambarkan keterampilan yang dimiliki siswa sebagai hasil dari proses pembelajaran. Untuk mendapatkan instrumen tes yang baik maka harus dilakukan analisis terhadap instrumen tersebut.

Analisis butir soal dalam penelitian ini menggunakan analisis model Rasch. Model Rasch dapat mengubah data skor mentah menjadi data dengan interval yang sama sehingga menghasilkan skala pengukuran yang linier, presisi dan mempunyai satuan (Sumintono & Widhiarso, 2015).

Pemodelan Rasch (Rasch Model Measurement) memiliki beberapa kelebihan dibandingkan dengan teori tes klasik. Kelebihan tersebut diantaranya, yaitu pertama mampu memberikan skala linier dengan interval yang sama; kedua, dapat melakukan prediksi terhadap data yang hilang; ketiga, bisa memberikan

estimasi yang lebih tepat; keempat, mampu mendeteksi ketidaktepatan model; dan kelima, menghasilkan pengukuran yang *replicable* (Sumintono & Widhiarso, 2014).

Melalui analisis model Rasch dengan bantuan aplikasi Ministep, karakteristik instrumen tes keterampilan berpikir kritis yang terdiri dari validitas, reliabilitas, tingkat kesukaran dan bias butir soal dapat diketahui.

Kajian Pustaka

1) Keterampilan Berpikir Kritis

Keterampilan berpikir kritis merupakan salah satu keterampilan yang harus dilatihkan kepada siswa pada abad ke-21 ini. Di banyak negara, berpikir kritis telah menjadi salah satu kompetensi dari tujuan pendidikan, bahkan sebagai salah satu sasaran yang ingin dicapai. Hal tersebut dilatarbelakangi kajian-kajian yang menunjukkan bahwa berpikir kritis merupakan keterampilan berpikir tingkat tinggi dan telah diketahui berperan dalam perkembangan moral, perkembangan sosial, perkembangan mental, perkembangan kognitif, dan perkembangan sains (Hashemi dkk, 2010).

Setiap orang memiliki keterampilan berpikir yang berbeda-beda. Untuk mengetahui keterampilan berpikir kritis yang dimiliki seseorang, kita dapat menggunakan indikator-indikatornya sebagai acuan. Facione (2018) membagi KBK menjadi beberapa indikator seperti pada Gambar 2.1, yaitu interpretasi (*interpretation*), analisis (*analysis*), kesimpulan (*inference*), evaluasi (*evaluation*), menjelaskan (*explanation*) dan pengaturan diri (*self-regulation*). Aspek-aspek tersebut telah dideskripsikan lebih lengkap dalam Facione (2018) sebagai berikut:

Tabel 1 Enam KBK Berdasarkan Facione

Aspek	Deskripsi	Sub-Aspek
Interpretasi	Memahami dan mengungkapkan makna dari berbagai pengalaman, situasi, data, kejadian, pertimbangan, konveksi (permufakatan atau kesepakatan), keyakinan, peraturan, prosedur atau kriteria.	(1) Mengategorikan (2) Menguraikan arti penting (3) Mengklarifikasi

Aspek	Deskripsi	Sub-Aspek
		makna
Analisis	Mengidentifikasi hubungan aktual yang dimaksud dan kesimpulan antara pernyataan, pertanyaan, konsep, deskripsi, atau bentuk representasi lainnya yang dimaksudkan untuk mengungkapkan keyakinan, penilaian, pengalaman, alasan, informasi, atau pendapat.	(1) Memeriksa gagasan (2) Mengidentifikasi argumen (3) Mengidentifikasi alasan dan klaim
Kesimpulan	Mengidentifikasi dan memilah-memilih hal yang dibutuhkan untuk menarik kesimpulan yang masuk akal; membuat dugaan dan hipotesis, mempertimbangkan informasi yang relevan dan mengurangi konsekuensi dari data yang diperoleh, pernyataan, prinsip, bukti pertimbangan, keyakinan, pendapat, konsep, deskripsi, pertanyaan, atau bentuk representasi lainnya.	(1) Meminta bukti (2) Membuat dugaan kesimpulan atau kesimpulan sementara (3) Membuat kesimpulan yang logis dan valid atas dasar pertimbangan
Evaluasi	Menilai kredibilitas pertanyaan atau representasi dalam bentuk lainnya seperti catatan, pengalaman, situasi, penilaian, keyakinan, atau opini seseorang; menilai hubungan kesimpulan antara pernyataan, deskripsi, pertanyaan, atau bentuk representasi lainnya.	(1) Menilai kredibilitas klaim (2) Menilai kualitas argumen induktif (dari khusus ke umum) atau deduktif.
Menjelaskan	Menyampaikan dan membenarkan pemikiran dalam hal-hal seperti bukti, konseptual, metodologis, kritik, dan pertimbangan kontekstual dari sebuah hasil pemikiran (penalaran) yang beralasan; menyampaikan pemikiran (penalaran) dalam bentuk argumen yang beralasan.	(1) Menyatakan hasil (2) Membenarkan atau memberikan alasan sebuah prosedur (3) Menyampaikan argumen
Pengaturan Diri	Secara sadar memantau diri dalam menerima pengetahuan, hal-hal yang digunakan dalam kegiatan pemantauan diri, dan membuat hasil (keputusan) terutama untuk menerapkan keterampilan analisis dan mengevaluasi sebuah kredibilitas simpulan dengan maksud mempertanyakan, mengkonfirmasi, memvalidasi atau mengoreksi penalaran diri atau orang lain.	(1) Pemantauan diri (2) Memeriksa diri

2) Model Rasch

Model Rasch pertama kali dikembangkan oleh matematikawan Denmark, Dr. Georg Rasch pada tahun 1950-an. Model Rasch memiliki kemiripan dengan model IRT (*Item Response Theory*) 1 parameter logistik (1PL). Model ini menghubungkan parameter orang (kemampuan siswa) dan parameter item (dalam hal ini tingkat kesukaran item) melalui suatu fungsi logaritma yang dapat menghasilkan pengukuran dengan interval yang sama, pengukuran ini memiliki satuan baru yang disebut logit (log odds unit). Dari nilai logit yang diperoleh, akan diketahui bahwa tingkat kesuksesan siswa dalam mengerjakan soal bergantung pada

tingkat kemampuan siswa dan tingkat kesukaran item (butir soal) (Olsen, dalam Sumintono, 2015).

2. Metode

Penelitian ini menggunakan metode kuantitatif deskriptif dengan desain konstruksi dan validasi yang merujuk pada desain yang dikembangkan oleh Crocker, L.M. & Algina, J. (1986). Dalam penelitian ini langkah-langkah yang harus dilakukan terdiri dari dua tahap utama yaitu tahap konstruksi yang terdiri dari proses 1) mengidentifikasi kegunaan tes; 2) mengidentifikasi karakteristik tes; 3) mempersiapkan spesifikasi tes; dan 4) mengkonstruksi butir soal, sedangkan

tahap validasi terdiri dari proses 1) review soal oleh ahli (judgement) dan revisi; 2) uji pendahuluan dan revisi; 3) uji lapangan; dan analisis tes dengan model rasch.

Proses review soal dilakukan oleh lima orang ahli, yang terdiri dari empat orang dosen fisika dan satu orang guru mata pelajaran fisika SMA. Kemudian hasil validasi dianalisis menggunakan teknik analisis CVR (*Content Validity Ratio*). Instrumen tes KBK yang sudah teruji validitasnya, kemudian diuji cobakan dalam skala kecil (uji pendahuluan) kepada 20 orang siswa SMA yang sudah mempelajari materi alat optik. Uji pendahuluan dilakukan tanpa ditentukan terlebih dahulu batas waktu pengerjaannya, hal tersebut dimaksudkan agar peneliti mendapatkan gambaran estimasi waktu pengerjaan yang dibutuhkan untuk uji lapangan. Pada uji pendahuluan pula dilakukan wawancara kepada lima orang siswa untuk mengetahui keterbacaan soal sebagai

bahan perbaikan. Setelah soal siap digunakan, kemudian dilakukan uji lapangan kepada 75 partisipan. Partisipan dalam penelitian ini adalah siswa siswi SMA sederajat yang telah mempelajari materi Alat Optik. Penelitian dilakukan di SMA yang ada di kota Bandung. Data berupa jawaban siswa yang dihasilkan dari uji lapangan kemudian dianalisis menggunakan model Rasch, untuk diketahui karakter tes nya seperti validitas tes, reliabilitas tes, tingkat kesukaran tes, dan adanya bias tes.

3. Hasil dan Pembahasan

Penelitian ini bermaksud untuk mengetahui karakteristik dari tes keterampilan berpikir kritis yang dikonstruksi. Hasil karakterisasi tes berdasarkan analisis model Rasch adalah sebagai berikut.

TABLE 10.1 D:\PERJALANAN SKRIPSI\7. LAMPIRAN\OLA ZOU273WS.TXTE Jul 24 2019 0:52 1\CAMPURAN
INPUT: 75 Person 24 Item REPORTED: 75 Person 24 Item 2 CATS MINISTEP 4.3.2

Person: REAL SEP.: 1.42 REL.: .67 ... Item: REAL SEP.: 3.60 REL.: .93

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	TOTAL MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
8	46	75	-.56	.25	1.05	.57	1.36	2.49	A .24	.36	74.7	67.9	S8
10	61	75	-1.70	.31	1.14	.78	1.33	1.13	B .10	.29	82.7	81.4	S10
6	18	75	1.31	.29	1.25	1.60	1.32	1.39	C .07	.34	68.0	77.6	S6
18	20	75	1.15	.28	1.16	1.16	1.26	1.26	D .17	.35	68.0	75.5	S18
4	54	75	-1.11	.27	1.20	1.53	1.24	1.21	E .12	.33	66.7	73.6	S4
3	26	75	.71	.26	1.03	.35	1.22	1.46	F .29	.36	76.0	69.9	S3
5	30	75	.45	.25	1.16	1.65	1.19	1.49	G .20	.37	57.3	67.3	S5
16	18	75	1.31	.29	.96	-.24	1.19	.87	H .32	.34	84.0	77.6	S16
19	34	75	.19	.25	1.06	.68	1.12	1.05	I .30	.37	73.3	66.0	S19
22	18	75	1.31	.29	.96	-.26	1.10	.50	J .34	.34	84.0	77.6	S22
1	60	75	-1.60	.30	.99	-.02	1.09	.43	K .28	.30	82.7	80.2	S1
23	39	75	-.12	.25	1.07	.84	1.03	.28	L .31	.37	58.7	65.7	S23
13	46	75	-.56	.25	.97	-.30	1.03	.24	l .38	.36	69.3	67.9	S13
21	38	75	-.05	.25	.97	-.31	.96	-.37	k .41	.37	65.3	65.6	S21
12	69	75	-2.74	.44	.93	-.10	.58	-.77	j .34	.21	92.0	92.0	S12
20	28	75	.58	.26	.91	-.95	.87	-.95	i .48	.37	73.3	68.5	S20
24	26	75	.71	.26	.91	-.86	.84	-1.11	h .48	.36	73.3	69.9	S24
17	47	75	-.62	.26	.90	-1.07	.90	-.72	g .46	.36	74.7	68.5	S17
15	44	75	-.43	.25	.89	-1.31	.84	-1.34	f .50	.37	72.0	67.0	S15
9	24	75	.85	.27	.88	-1.09	.79	-1.35	e .51	.36	73.3	71.8	S9
2	54	75	-1.11	.27	.87	-1.01	.76	-1.28	d .49	.33	74.7	73.6	S2
11	32	75	.32	.25	.87	-1.47	.85	-1.33	c .51	.37	76.0	66.6	S11
14	25	75	.78	.26	.87	-1.25	.80	-1.36	b .52	.36	72.0	70.9	S14
7	23	75	.92	.27	.85	-1.33	.82	-1.08	a .52	.36	76.0	72.7	S7
MEAN	36.7	75.0	.00	.27	.99	-.1	1.02	.1			73.7	72.3	
P. SD	14.9	.0	1.06	.04	.12	1.0	.21	1.2			7.7	6.2	

Gambar 1. Validitas Tes KBK

a. Validitas Tes

Pada Gambar 1 jika kita tinjau kolom *entry number* menunjukkan nomor soal yang diurutkan dari soal yang memiliki validitas rendah menuju soal yang memiliki validitas tinggi sedangkan kolom *outfit* dan kolom *ptmeasure* menunjukkan kriteria validitas.

Tes yang valid memenuhi minimal dua dari tiga kriteria validitas. Kriteria tersebut yaitu *outfit means-square* (MNSQ), *outfit z-standard* (ZSTD) dan *point measure correlation* (CORR.).

Pada Gambar 1 diatas terlihat bahwa semua butir soal memenuhi kriteria MNSQ dengan nilai MNSQ tertinggi 1,36 dan nilai terendah 0,58. Pada kriteria kedua yaitu ZSTD, hanya satu butir yang tidak memenuhi kriteria tersebut yaitu butir

soal S8 dengan nilai ZSTD sebesar +2,49. Kemudian pada kriteria ketiga yaitu point measure correlation yang disingkat CORR terdapat 14 butir soal yang tidak memenuhi kriteria diantaranya butir soal S1, S3, S4, S5, S6, S8, S10, S12, S13, S16, S18, S19, S22 dan S23. Jadi dapat dikatakan bahwa butir soal yang paling atas yaitu S8 mempunyai kecenderungan tidak fit atau tidak valid, karena tidak memenuhi dua dari tiga kriteria validitas. Sedangkan untuk butir soal yang lainnya hanya tidak memenuhi satu kriteria saja. Maka dapat disimpulkan bahwa pada hasil uji validitas tes KBK terdapat satu soal yang dikategorikan tidak valid yaitu butir soal S8.

TABLE 3.1 D:\PERJALANAN SKRIPSI\7. LAMPIRAN\OLAH ZOU273WS.TXTP Jul 24 2019 0:521\CAMPURAN INPUT: 75 Person 24 Item REPORTED: 75 Person 24 Item 2 CATS MINISTEP 4.3.2

SUMMARY OF 75 MEASURED Person									
	TOTAL SCORE	COUNT	MEASURE	MODEL S. E.	INFIT		OUTFIT		
					MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	11.7	24.0	-.02	.48	.99	-.03	1.02	.06	
SEM	.4	.0	.10	.00	.02	.10	.03	.09	
P. SD	3.8	.0	.86	.03	.17	.84	.25	.81	
S. SD	3.9	.0	.87	.03	.17	.84	.25	.81	
MAX.	20.0	24.0	1.92	.60	1.60	2.29	1.94	2.50	
MIN.	4.0	24.0	-1.95	.45	.64	-2.01	.57	-1.95	
REAL RMSE	.50	TRUE SD	.71	SEPARATION	1.42	Person RELIABILITY		.67	
MODEL RMSE	.48	TRUE SD	.72	SEPARATION	1.49	Person RELIABILITY		.69	
S. E. OF Person MEAN = .10									
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00									
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .69 SEM = 2.15									
SUMMARY OF 24 MEASURED Item									
	TOTAL SCORE	COUNT	MEASURE	MODEL S. E.	INFIT		OUTFIT		
					MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	36.7	75.0	.00	.27	.99	-.10	1.02	.09	
SEM	3.1	.0	.22	.01	.02	.21	.04	.24	
P. SD	14.9	.0	1.06	.04	.12	.99	.21	1.16	
S. SD	15.2	.0	1.08	.04	.12	1.01	.22	1.19	
MAX.	69.0	75.0	1.31	.44	1.25	1.65	1.36	2.49	
MIN.	18.0	75.0	-2.74	.25	.85	-1.47	.58	-1.36	
REAL RMSE	.28	TRUE SD	1.02	SEPARATION	3.60	Item RELIABILITY		.93	
MODEL RMSE	.28	TRUE SD	1.02	SEPARATION	3.69	Item RELIABILITY		.93	
S. E. OF Item MEAN = .22									

Gambar 2. Reliabilitas Tes KBK

b. Reliabilitas Tes

Pada Gambar 2 tinjau nilai *person reliability* yang menunjukkan tingkat reliabilitas siswa, nilai *item reliability* yang menunjukkan tingkat reliabilitas item/butir soal, dan nilai *cronbach alpha* yang menunjukkan nilai reliabilitas instrumen secara keseluruhan.

Berdasarkan Gambar 2 diatas dapat diketahui bahwa instrumen tes KBK yang telah dikonstruksi memiliki nilai *person reliability* sebesar 0,67 dan *item reliability* sebesar 0,93. Nilai tersebut memberikan informasi bahwa siswa cukup konsisten dalam menjawab soal-soal yang diujikan,

dan juga kualitas butir-butir soal dalam instrumen tes KBK memiliki reliabilitas yang bagus sekali. Sehingga interaksi antara *person* dan *item* bisa dilihat pada nilai cronbach alpha yaitu sebesar 0,69 yang menunjukkan bahwa secara keseluruhan instrumen memiliki reliabilitas yang cukup.

Untuk dapat mengetahui aspek KBK dan butir soal mana yang memiliki nilai reliabilitas dengan kategori lemah sampai istimewa maka dilakukan uji reliabilitas testiap sub-aspek KBK. Hasil Uji reliabilitas disajikan dalam Tabel 1.

Tabel 2. Hasil Uji Reliabilitas Tes Per Sub-Aspek KBK

Aspek	Sub-aspek	Soal No-	Jenis Reliabilitas	Nilai Reliabilitas	Keterangan
Interpretasi	1	1, 2	<i>Item Reliability</i>	0,79	Cukup
	2	3, 4	<i>Item Reliability</i>	0,93	Bagus Sekali
	3	5, 6	<i>Item Reliability</i>	0,71	Cukup
Analisis	1	7, 8	<i>Item Reliability</i>	0,91	Bagus Sekali
	2	9	<i>Item Reliability</i>	0,00	Lemah
	3	10, 11	<i>Item Reliability</i>	0,94	Bagus Sekali
Kesimpulan	1	12, 13	<i>Item Reliability</i>	0,92	Bagus Sekali
	2	14	<i>Item Reliability</i>	0,00	Lemah
	3	15, 16	<i>Item Reliability</i>	0,93	Bagus Sekali
Evaluasi	1	17, 18	<i>Item Reliability</i>	0,93	Bagus Sekali
	2	19	<i>Item Reliability</i>	0,00	Lemah
Menjelaskan	1	20, 21	<i>Item Reliability</i>	0,74	Cukup
	2	22	<i>Item Reliability</i>	0,00	Lemah
	3	23, 24	<i>Item Reliability</i>	0,86	Bagus

Dari Tabel 1 terlihat bahwa berdasarkan hasil uji reliabilitas tes per sub-aspek KBK pada kelompok pertama terdapat 6 sub-aspek yang termasuk kategori bagus sekali (soal no. 3, 4, 7, 8, 10, 11, 12, 13, 15, 16, 17, dan 18), 1 sub

aspek kategori bagus (soal no. 23 dan 24), 3 sub-aspek kategori cukup (soal no. 1, 2, 5, 6, 20, dan 21), dan 4 sub-aspek kategori lemah (soal no. 9, 14, 19, dan 22).

INPUT: 75 Person 24 Item REPORTED: 75 Person 24 Item 2 CATS MINISTEP 4.3.2
 Person: REAL SEP.: 1.42 REL.: .67 ... Item: REAL SEP.: 3.60 REL.: .93

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S. E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
6	18	75	1.31	.29	1.25	1.60	1.32	1.39	.07	.34	68.0	77.0	S6
16	18	75	1.31	.29	.96	-.24	1.19	.87	.32	.34	84.0	77.0	S16
22	18	75	1.31	.29	.96	-.26	1.10	.50	.34	.34	84.0	77.0	S22
18	20	75	1.15	.28	1.16	1.16	1.26	1.26	.17	.35	68.0	75.0	S18
7	23	75	.92	.27	.85	-1.33	.82	-1.08	.52	.36	76.0	72.0	S7
14	24	75	.85	.27	.88	-1.09	.79	-1.35	.51	.36	73.3	71.0	S9
14	25	75	.78	.26	.87	-1.25	.80	-1.36	.52	.36	72.0	70.0	S14
3	26	75	.71	.26	1.03	-.35	1.22	1.46	.29	.36	76.0	69.0	S3
24	26	75	.71	.26	.91	-.86	.84	-1.11	.48	.36	73.3	69.0	S24
20	28	75	.58	.26	.91	-.95	.87	-.95	.48	.37	73.3	68.0	S20
5	30	75	.45	.25	1.16	1.65	1.19	1.49	.20	.37	57.3	67.0	S5
11	32	75	.32	.25	.87	-1.47	.85	-1.33	.51	.37	76.0	66.0	S11
19	34	75	-.19	.25	1.06	.68	1.12	1.05	.30	.37	73.3	66.0	S19
21	38	75	-.05	.25	.97	-.31	.96	-.37	.41	.37	65.3	65.0	S21
23	39	75	-.12	.25	1.07	.84	1.03	.28	.31	.37	58.7	65.0	S23
15	44	75	-.43	.25	.89	-1.31	.84	-1.34	.50	.37	72.0	67.0	S15
8	46	75	-.56	.25	1.05	.57	1.36	2.49	.24	.36	74.7	67.0	S8
13	46	75	-.56	.25	.97	-.30	1.03	.24	.38	.36	69.3	67.0	S13
17	47	75	-.62	.26	.90	-1.07	.90	-.72	.46	.36	74.7	68.0	S17
2	54	75	-1.11	.27	.87	-1.01	.76	-1.28	.49	.33	74.7	73.0	S2
4	54	75	-1.11	.27	1.20	1.53	1.24	1.21	.12	.33	66.7	73.0	S4
1	60	75	-1.60	.30	.99	-.02	1.09	.43	.28	.30	82.7	80.0	S1
10	61	75	-1.70	.31	1.14	.78	1.33	1.13	.10	.29	82.7	81.0	S10
12	69	75	-2.74	.44	.93	-.10	.58	-.77	.34	.21	92.0	92.0	S12
MEAN	36.7	75.0	.00	.27	.99	-.1	1.02	.1			73.7	72.0	
P. SD	14.9	.0	1.06	.04	.12	1.0	.21	1.2			7.7	6.0	

Gambar 3. Tingkat Kesukaran Tes

c. Tingkat Kesukaran

Pada Gambar 3 kolom *measure* menunjukkan nilai logit yang dimiliki setiap soal dari terbesar ke terkecil sedangkan kolom *item* menunjukkan urutan nomor soal dari yang tersukar ke yang termudah.

Berdasarkan Gambar 3 diatas dapat diketahui bahwa butir soal S6, S16 dan S22 memiliki tingkat kesukaran yang paling tinggi diantara butir soal yang lainnya dengan nilai logit masing-masing sebesar +1,31 (kolom *measure*) dan terdapat 18 orang yang dapat menjawab

dengan benar butir soal tersebut (kolom total score) dari 75 orang peserta tes. Sedangkan untuk butir soal yang memiliki tingkat kesukaran paling rendah (mudah) adalah butir soal S12 dengan nilai logit sebesar -2,74 dan hampir semua peserta tes dapat menjawab soal tersebut dengan benar kecuali 6 peserta tes.

Dalam tabel diatas diketahui bahwa nilai standar deviasinya 1,06. Nilai ini dapat dijadikan acuan jika ingin mengelompokkan tingkat kesukaran soalnya.

Tabel 3. Pengelompokan Tingkat Kesukaran Butir Soal Tes KBK

Kelompok Soal	Acuan	Butir Soal
Sangat Sukar	<i>Logit</i> > 1,06	S6, S16, S22, S18
Sukar	(0,0 <i>logit</i> + 1SD) <i>Logit</i> (0,0 - (+1,06))	S3, S5, S7, S9, S11, S14, S20, S24
Mudah	(0,0 <i>logit</i> - 1SD) <i>Logit</i> (0,0 - (-1,06))	S8, S13, S15, S17, S21, S23
Sangat Mudah	<i>Logit</i> > - 1,06	S1, S2, S4, S10, S12

TABLE 30.4 D:\PERJALANAN SKRIPSI\7. LAMPIRAN\OLA ZOU273WS.TXTE Jul 24 2019 0:52 1\CAMPURAN
 INPUT: 75 Person 24 Item REPORTED: 75 Person 24 Item 2 CATS MINISTEP 4.3.2

DIF class/group specification is: DIF=\$S3W1

Person CLASSES	SUMMARY DIF CHI-SQUARED	D. F.	PROB.	BETWEEN-CLASS/GROUP UNWTD MNSQ	ZSTD	Item Number	Name
2	.0069	1	.9337	.0089	-1.21	1	S1
2	5.8981	1	.0152	6.5685	2.32	2	S2
2	.0304	1	.8617	.0310	-.98	3	S3
2	2.0210	1	.1551	2.1063	1.07	4	S4
2	2.4410	1	.1182	2.5488	1.25	5	S5
2	.0242	1	.8765	.0247	-1.03	6	S6
2	1.1503	1	.2835	1.1824	.59	7	S7
2	3.8901	1	.0486	4.1398	1.76	8	S8
2	.6825	1	.4087	.6977	.23	9	S9
2	9.3528	1	.0022	11.0642	3.08	10	S10
2	.2594	1	.6106	.2618	-.29	11	S11
2	1.7280	1	.1887	1.8153	.94	12	S12
2	.9804	1	.3221	1.0064	.48	13	S13
2	.3388	1	.5605	.3472	-.16	14	S14
2	1.2362	1	.2662	1.2707	.65	15	S15
2	4.3071	1	.0380	4.6634	1.89	16	S16
2	.5730	1	.4491	.5869	.13	17	S17
2	8.0431	1	.0046	9.0668	2.77	18	S18
2	8.4607	1	.0036	9.4909	2.84	19	S19
2	.1569	1	.6920	.1580	-.50	20	S20
2	1.3701	1	.2418	1.4130	.73	21	S21
2	8.9928	1	.0027	11.0079	3.07	22	S22
2	8.3113	1	.0039	9.3216	2.81	23	S23
2	1.4546	1	.2278	1.5016	.78	24	S24

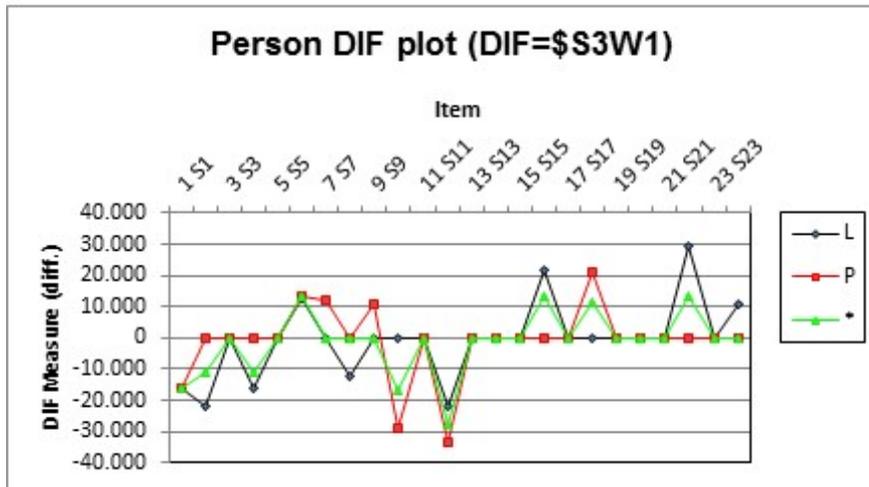
Gambar 4. Bias Butir Tes KBK

d. Bias Butir

Pada Gambar 4 kolom *prob* menunjukkan nilai probabilitas setiap soal sedangkan kolom *name* menunjukkan nomor soal.

Berdasarkan Gambar 4 di atas dapat diketahui bahwa pada butir soal S2, S8,

S10, S16, S18, S19, S22 dan S23 memiliki nilai probabilitas kurang dari 5%. Hal tersebut menunjukkan bahwa soal-soal tersebut mengandung bias sehingga perlu diperbaiki agar tidak merugikan kelompok gender tertentu.



Gambar 5. Grafik Deteksi Bias Butir Tes KBK

Berdasarkan Gambar 5 diatas terlihat bahwa butir soal S2 dan S18 mudah dikerjakan oleh laki-laki (ada di bawah), dibandingkan siswa perempuan. Sedangkan untuk butir soal S10, S16 dan S22 terlihat kurva warna merah berada lebih kebawah, hal tersebut menandakan bahwa soal tersebut mudah dikerjakan oleh perempuan.

Dari hasil penelitian yang didapatkan terlihat bahwa terdapat satu soal yang tidak valid yaitu soal S8. Hal tersebut bisa disebabkan karena siswa kurang termotivasi saat mengerjakan tes menyebabkan banyak siswa yang asal-asalan mengerjakan tes. Dampaknya adalah respons mereka bervariasi dan tidak konsisten, tergantung pada mood mereka saat merespons butir (Fisher, 1993). Faktor lainnya bisa disebabkan karena ukuran sampel yang dijadikan partisipan kurang begitu banyak. Jumlah partisipan hanya terdiri dari 75 siswa. Hal tersebut sejalan dengan pernyataan Sumintono dan Widhiarso (2015) yang mengatakan bahwa nilai kesesuaian butir (validitas) sangat dipengaruhi besarnya ukuran sampel. Ia juga mengatakan faktor yang menjadikan soal tidak valid diantaranya, banyaknya individu yang kurang termotivasi dalam mengerjakan soal, atau soal yang memiliki daya beda rendah.

Dalam pemodelan Rasch, reliabilitas yang dihasilkan merupakan reliabilitas separasi. Reliabilitas ini melaporkan 2 hal yaitu yaitu reliabilitas butir (*item reliability*) dan reliabilitas orang (*person reliability*). Dalam penelitian ini terdapat beberapa soal yang memiliki nilai reliabilitas yang rendah sehingga dikategorikan lemah yaitu soal no. 9, 14, 19, dan 22. Biasanya reliabilitas yang rendah dikarenakan sampel terlalu sedikit sehingga variasi hirarki pada penggaris logit hanya sedikit (Linacre, 2016 dalam Akhtar, 2017).

Hasil penelitian menunjukkan bahwa pada tes yang dikonstruksi terdapat beberapa soal yang mengandung bias terhadap jenis kelamin yaitu soal S2, S8, S10, S16, S18, S19, S22 dan S23 karena memiliki nilai probabilitas kurang dari 5%. Suatu butir soal disebut bias jika didapati bahwa individu dengan karakteristik

tertentu lebih diuntungkan dalam menjawab soal dibanding individu dengan karakteristik lain (Sumintono & Widhiarso, 2015). Faktor yang mempengaruhi terjadinya bias butir pada instrumen tes adalah: 1) faktor peserta didik itu sendiri, biasanya diakibatkan karena kondisi psikologis dan kesehatan pada saat mengikuti tes; 2) faktor dari proses pembelajaran, faktor ini terjadi ketika anak didik dalam memaknai ucapan atau penjelasan dari guru yang kurang memahami materi yang disajikan oleh guru; 3) faktor dari pelaksana tes, karena pengawas pada saat itu tidak jeli atau kurang mengawasi pada saat tes berlangsung sehingga peserta didik bebas menyontek pada buku catatan atau sesama peserta didik yang lain (peserta tes); 4) faktor lingkungan, terjadi karena suasana lingkungan pada saat itu kurang mendukung misalnya sekitar ruangan ujian yang terlalu ribut oleh kelas lain yang tidak belajar sehingga mengganggu konsentrasi peserta tes (Nurhidayah, 2016).

4. Simpulan

Tes keterampilan berpikir kritis materi alat optik terdiri dari 24 soal pilihan ganda (*multiple choice*) yang mengacu pada 5 dari 6 aspek KBK menurut Facione (2018) yaitu: interpretasi, analisis, kesimpulan, evaluasi dan menjelaskan. Hasil analisis data berdasarkan model Rasch menunjukkan: 1) Validitas tes keterampilan berpikir kritis pada materi alat optik dalam penelitian ini diketahui memiliki validitas isi tes yang valid berdasarkan analisis CVR (*Content Validity Ratio*). Kemudian, validitas empiris yang dihasilkan pada uji lapangan diketahui 23 soal dari 24 soal dikategorikan valid. Soal yang tidak valid yaitu soal S8 karena tidak memenuhi minimal 2 dari 3 kriteria validitas (MNSQ, ZSTD, dan *point measure correlation*). 2) Reliabilitas tes keterampilan berpikir kritis pada materi alat optik berdasarkan analisis model Rasch dalam penelitian ini termasuk kategori cukup. Dengan nilai cronbach alpha (reliabilitas tes) sebesar 0,69. Serta memiliki nilai item reliability

sebesar 0,93 (kategori bagus sekali) dan person reliability sebesar 0,67 (kategori cukup). Beberapa soal memiliki reliabilitas item yang lemah yaitu soal S9, S14, S19, dan S22. 3) Tingkat kesukaran dalam penelitian ini menunjukkan bahwa soal S6, S16 dan S22 memiliki tingkat kesukaran yang paling tinggi karena memiliki nilai logit paling besar yaitu +1,31 dan soal S12 memiliki tingkat kesukaran yang paling rendah karena memiliki nilai logit -2,74. 4) Hasil deteksi adanya bias butir soal tes keterampilan berpikir kritis berdasarkan analisis model Rasch dalam penelitian ini diketahui berdasarkan nilai probabilitas yang kurang dari 0,05. Soal yang mengandung bias terdapat pada soal S2 dan S18 karena lebih menguntungkan siswa laki-laki dan soal S10, S16 dan S22 lebih menguntungkan siswa perempuan.

Daftar Pustaka

- Akhtar, Hanif. 2017. *Berkenalan Dengan Rasch Model*. [Online]. Diakses dari <https://www.semestapsikometrika.com/2017/07/berkenalan-dengan-rasch-model.html>.
- Facione, P. A. 2018. Critical thinking: What it is and why it counts. *Insight assessment*, 2007(1), 1-23.
- Fisher, R. J. 1993. Social desirability bias and the validity of indirect questioning. *Journal of consumer research*, 20(2), 303-315.
- Fisher, A. 2009. *Berpikir Kritis: Sebuah Pengantar*. Diterjemahkan oleh: Hadinata, B. Jakarta: Erlangga.
- Hashemi, SA, dkk. 2010. Science Production In Iranian Educational System By The Use Of Critical Thinking. *International Journal of Instruction* January 2010. Vol.3, No.1
- Oonsim, W., dan Chanprasert, K. 2017. Developing Critical Thinking Skills of Grade 11 Students by STEM Education: A Focus on Electrostatic in Physics. *Rangsit Journal of Educational Studies*, Vol.4, No.1, pp.54-59.
- Ozsevgec, T., dan Cepni, S. 2006. Relation between Science Teachers' Assessment Tools and Students Cognitive Development. *Online Submission*, 1(7), 222-226.
- Sugiarti, T., dkk. 2017. Development of Assessment Instrument of Critical Thinking in Physics at Senior High School. In *Journal of Physics:Conference Series* (Vol. 812, No. 1, p. 012018). IOP Publishing.
- Sumintono, B., dan Widhiarso, W. 2014. *Aplikasi model Rasch untuk penelitian ilmu-ilmu sosial (edisi revisi)*. Trim Komunikata Publishing House.
- Sumintono, B., dan Widhiarso, W. 2015. *Aplikasi pemodelan rasch pada assessment pendidikan*. Trim Komunikata.
- Wagner, T. 2010. *Overcoming the global achievement gap*. Cambridge, Mass., Harvard University.
- Wartono, W., dkk. 2018. How are the physics critical thinking skills of the students taught by using inquiry-discovery through empirical and theoretical overview. *Eurasia Journal of Mathematics, Science and Technology Education*, 14(2), 691-697.
- Zubaidah, S. 2010. Berpikir Kritis: Kemampuan Berpikir Tingkat Tinggi yang Dapat Dikembangkan melalui Pembelajaran Sains. In *Makalah Seminar Nasional Sains dengan Tema Optimalisasi Sains untuk memberdayakan Manusia*. Pascasarjana Unesa (Vol. 16).