

Analisis Tes Keterampilan Berpikir Kritis Remaja (TKBKR) Pada Mahasiswa Pendidikan Biologi : Analisis Pemodelan Rasch

The Analysis of Critical Thinking Skills Test of Adolescent (TKBKR) in Biology Education Students : A Rasch Model Analysis

Lucky Nindi R. Marfu'i¹, Ilfiandra² dan Nurhudaya³

¹²³ Universitas Pendidikan Indonesia

Departemen Psikologi Pendidikan dan Bimbingan, Bandung, Indonesia

riandika.luckybk@student.upi.edu

Naskah diterima 26/07/2018; direvisi 28/07/2018; disetujui 01/08/2018

Abstract. Construction of test equipment requires a trial before the test is issued. It should be noted that before the trial is needed an judgment to the expert. Tests of constructing critical thinking skills are generic. This test contains several questions that are described by the problems of teenagers' lives in terms of economic, social, cultural, and educational. Based on the statement, it is necessary to test the external validity to students with a scientific background. The method used in this study is descriptive evaluation with a quantitative approach. The data analysis used in this study is the RASCH modeling analysis. The reliability coefficient in the RASCH modeling analysis is divided into three which include; 1) overall reliability, 2) reliability of items, and 3) reliability of respondents in answering questions. The results of this study are a reference in the development and revision of tests of critical thinking skills, as well as the results of the first trial. In this study, the TKBKR in the first external validity test showed reliability in the moderate category with an overall reliability coefficient of 0.65. This instrument contains 36 questions with the question that including the maximum extreme score (questions that are very difficult and inaccessible to respondents) only one question. This instrument can be used in scientific fields with a note of revising the product by decreasing the difficulty level of the question and revising the item.

Keywords: Critical Thinking Skills Test of Adolescent (TKBKR), Biology Education, Test Cognitive Construction, RASCH Model Analysis.

Abstrak. Konstruksi alat tes memerlukan uji coba sebelum tes diterbitkan. Perlu diketahui bahwa sebelum uji coba diperlukan adanya *judgement* instrumen kepada ahli. Tes keterampilan berpikir kritis yang dikonstruksi bersifat generik. Tes ini memuat beberapa soal yang digambarkan dengan problematika kehidupan remaja dari segi ekonomi, sosial, budaya, dan pendidikan. Berdasarkan pernyataan tersebut, perlu adanya pengujian validitas eksternal kepada mahasiswa berlatar-belakang keilmuan saintifik. Metode yang digunakan dalam penelitian ini adalah evaluasi deskriptif dengan pendekatan kuantitatif. Analisis data yang digunakan dalam penelitian ini adalah analisis pemodelan RASCH. Koefisien reliabilitas dalam analisis pemodelan RASCH dibagi menjadi tiga yang meliputi; 1) reliabilitas keseluruhan, 2) reliabilitas item, dan 3) reliabilitas responden dalam menjawab soal. Hasil penelitian ini menjadi acuan dalam pengembangan dan bahan revisi tes keterampilan berpikir kritis, serta sebagai hasil uji coba pertama. Dalam penelitian ini menggambarkan bahwa TKBKR pada uji validitas eksternal pertama menunjukkan reliabilitas pada kategori moderat dengan koefisien reliabilitas keseluruhan sebesar 0.65. Instrumen ini memuat 36 soal dengan soal yang termasuk

maximum extreme score (soal yang sangat sulit dan tidak dapat dijangkau responden) hanya satu soal. Instrumen ini dapat dipergunakan untuk bidang keilmuan saintifik dengan catatan merevisi produk dengan melakukan penurunan tingkat kesulitan soal dan merevisi item.

Kata kunci: Tes Keterampilan Berpikir Kritis Remaja, Pendidikan Biologi, Konstruksi Tes Kognitif, Analisis Pemodelan RASCH.

PENDAHULUAN

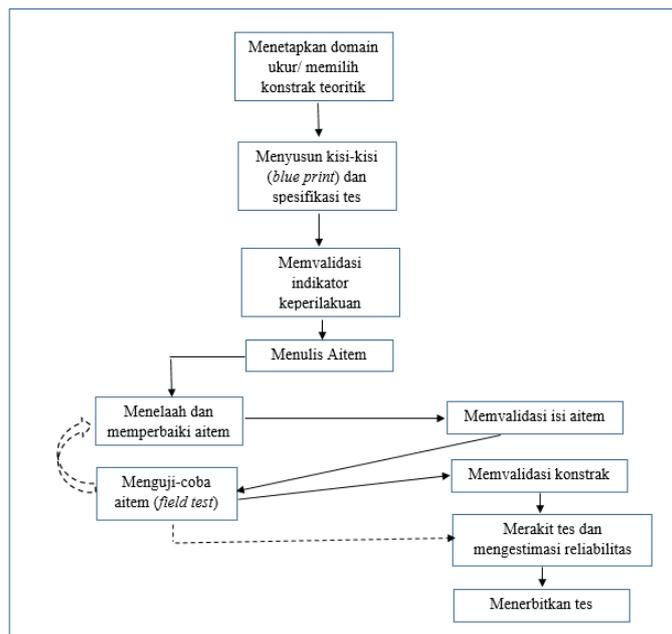
Keterampilan berpikir kritis merupakan salah satu inti dari pendidikan di abad 21. Keterampilan berpikir kritis merupakan proses berpikir secara mendalam dan sistematis mengenai suatu informasi atau permasalahan yang menuntut individu untuk mengambil keputusan dalam memecahkan masalah yang dialami. Artinya, keterampilan berpikir kritis bersifat kompleks dan tidak hanya membahas mengenai satu keterampilan saja dalam diri individu.

Dalam artikel ini akan membahas tiga pembahasan inti yaitu meliputi;

- 1) konstruksi tes keterampilan kognitif,
- 2) keterampilan berpikir kritis mahasiswa, dan
- 3) analisis pemodelan RASCH yang akan dijelaskan lebih lanjut sebagai berikut.

Konstruksi tes keterampilan kognitif

Tes ini dikonstruksi berdasarkan langkah-langkah pengembangan instrumen tes kognitif dalam asesmen pendidikan. Berikut adalah langkah-langkah yang harus ditempuh dalam konstruksi instrumen berikut.



Gambar 1. Langkah-langkah Konstruksi Tes (Saifuddin Azwar, 2016)

Pada gambar di atas dapat diketahui langkah pembuatan tes yang digunakan oleh peneliti dimulai dari menentukan domain alat ukur atau konstruk alat ukur. Dalam penelitian ini sudah spesifik ditentukan bahwa yang akan dianalisis adalah tes keterampilan kognitif dengan spesifikasi tes keterampilan berpikir kritis.

Aspek-aspek yang diukur menggunakan tes ini meliputi enam aspek yaitu 1) interpretasi, 2) analisis, 3) evaluasi, 4) inferensi, 5) menjelaskan, dan 6) regulasi diri. Komponen dari

keenam aspek ini dijelaskan pada tabel 1. Selanjutnya adalah menyusun kisi-kisi instrumen dengan menggunakan acuan keenam aspek tersebut menjadi komponen dalam menentukan dan menuliskan soal.

Langkah berikutnya adalah melakukan validasi indikator dengan ahli (*expert judgement*). Ahli yang ditunjuk untuk memvalidasi konten ada tiga, yakni dosen dari program studi fisika spesifik konstruksi tes keterampilan berpikir kritis, kedua adalah dosen dari pendidikan bahasa dan

sastra Indonesia yang memvalidasi bahasa yang digunakan dalam tes agar tidak terjadi miskonsepsi, dan ketiga adalah dosen dari program studi pendidikan matematika untuk memeriksa kesesuaian dan seberapa analitis kunci jawaban terhadap soal yang diproyeksikan.

Setelah itu dilanjutkan dengan menulis, menelaah, dan memperbaiki item. Setelah dilakukan validasi konten oleh para ahli, instrumen ditulis dan diperbaiki sesuai saran ahli sehingga siap untuk diuji-cobakan ke mahasiswa pendidikan biologi sebagai salah satu kelompok sampel mahasiswa yang memiliki latar belakang keilmuan saintifik.

Keterampilan Berpikir Kritis dalam Pendidikan

Mahasiswa pendidikan biologi memiliki karakteristik bermacam-macam. Salah satunya adalah mereka seringkali dihadapkan pada praktikum, sehingga waktu untuk bersosialisasi dan mengalami permasalahan dalam interaksi sosial sangat minim. Selain itu, mereka terbiasa berpikir secara analitis sehingga dimungkinkan dalam mengerjakan uji coba tes keterampilan berpikir kritis ini responden akan sungguh-sungguh teliti.

Bidang keilmuan saintifik sering dikenal dengan kemampuan kognitifnya yang menonjol. Dalam penelitian ini, penulis akan mengambil sampel untuk pengembangan instrumen tes kepada mahasiswa yang berlatar belakang saintifik salah satunya mahasiswa program studi pendidikan biologi karena dianggap akan dikerjakan dengan sungguh-sungguh berpikir kritis.

Aspek-aspek berpikir kritis dapat dijelaskan pada tabel berikut.

Tabel 1. Aspek-aspek Berpikir Kritis

No	Aspek	Indikator
1	Interpretasi (<i>Interpretation</i>)	a. Mengkategorisasi b. Mengartikan konten c. Mengklarifikasi makna secara implisit maupun eksplisit
2	Analisis (<i>Analysis</i>)	a. Menguji ide-ide. b. Mengidentifikasi argumen. c. Menganalisa argumen
3	Evaluasi (<i>Evaluation</i>)	a. Menilai kredibilitas informasi atau opini b. Menilai kualitas argumen menggunakan penalaran induktif dan deduktif

4	Inferensi (<i>Inference</i>)	a. Mengenali bukti-bukti atau fakta dari sebuah informasi b. Menyusun hipotesis alternatif c. Menjelaskan kesimpulan menggunakan penalaran induktif dan deduktif
5	Penjelasan (<i>Explanation</i>)	a. Menyatakan hasil. b. Menyesuaikan prosedur. c. Mempresentasikan argumen.
6	Regulasi diri (<i>Self regulation</i>)	a. Memonitor diri b. Mengoreksi diri

Source: *Facione, 1990; Marfu'i, 2018*

Berpikir kritis merupakan suatu landasan fundamental untuk memahami isu atau problematika yang bersifat kompleks (Heinrich, Habron, Johnson, & Goralnik, 2015). Berpikir kritis sangat disarankan untuk diukur menggunakan pendekatan saintifik. Hal ini dikarenakan berpikir kritis merupakan konstruk keterampilan yang kompleks, penempatan, dan metakognisi yang menunjukkan beberapa tantangan bagi mereka yang ingin mengassess berpikir kritis secara komprehensif (Bensley & Murtagh, 2012b, 2012a). Pendapat lain menambahkan, berpikir kritis juga dapat menautkan antara pendidikan dengan permasalahan kewarganeraan (Celuch, 2009). Berbagai hal, permasalahan, isu, dan informasi yang bersifat kompleks dapat diselesaikan menggunakan berpikir kritis.

Keterampilan berpikir kritis berkorelasi dengan persepsi mengenai keterampilan individu dalam memecahkan permasalahan. Persaingan di era global menuntut kemampuan berpikir dalam pengambilan keputusan sebagai pondasi praktik pendidikan (Garrett, 2013). Urgensi keterampilan berpikir kritis keterampilan berpikir kritis dengan keterampilan pemecahan masalah mendukung keterampilan *melek* teknologi informasi (Chaeruman, dalam Suarsana, 2013). Kemampuan berpikir pada individu sangat sulit untuk dikembangkan, akan tetapi hal itu masih mungkin untuk diupayakan (Kargar, et.al., 2013). Berpikir kritis adalah aktivitas dengan pemikiran yang mendalam dalam memenuhi beragam standar intelektual seperti kejelasan, relevansi, kecakapan, koherensi dan lain-lain (Fisher, 2009). dalam persaingan di era globalisasi menuntut individu untuk lebih kritis dalam menyikapi permasalahan di beberapa bidang. Keterampilan berpikir kritis diperlukan dalam menyikapi dan

menelaah beberapa informasi atau sumber berita di era globalisasi. Berdasarkan pendapat-pendapat tersebut, dapat disimpulkan bahwa keterampilan berpikir kritis perlu diukur untuk mengetahui kemampuan kognisi individu.

Berpikir kritis juga digunakan untuk menyelesaikan hal-hal yang bersifat abstrak, terutama dalam memecahkan permasalahan sebagai penunjang untuk menyempurnakan tujuan pembelajaran dan pengembangan individu (Miri, 2007). Berdasarkan pendapat di atas, maka diperlukan pengukuran terhadap keterampilan berpikir kritis pada remaja karena keterampilan ini bersifat kompleks dan memerlukan pengukuran secara komprehensif.

Analisis Pemodelan RASCH

Dalam analisis pengukuran terdapat teori modern dan teori klasik. Analisis pemodelan RASCH merupakan analisis menggunakan teori modern. Seperti yang diketahui bahwa analisis pemodelan RASCH baru sering digunakan sekitar tahun 2015. Analisis RASCH memiliki beberapa kelebihan, salah satunya dalam analisis RASCH tidak hanya memperhatikan skor mentah saja dalam menilai abilitas responden akan tetapi juga memperhatikan pola jawaban responden dan tingkat kesulitan soal (Marfu'i, 2018b).

Dalam analisis ini dapat digunakan untuk memprediksi abilitas responden terhadap kemampuan yang hendak diukur. Analisis ini menjelaskan bahwa, *“seseorang dengan yang memiliki abilitas tinggi akan memiliki probabilitas yang lebih besar untuk menjawab soal dengan benar. Hal yang sama juga berlaku untuk butir. Butir yang memiliki tingkat kesulitan lebih tinggi memiliki probabilitas untuk menyelesaikan butir tersebut lebih rendah daripada butir yang lain”*. (Rasch, Kubinger, & Yanagida, 2011). Pernyataan ini jika disimpulkan dengan singkat yakni dalam analisis Rasch skor yang tinggi tidak mencerminkan kemampuan seseorang individu yang tinggi pula, akan tetapi perlu ditelusuri kembali soal dengan tingkat kesulitan seperti apa dan seberapa besar probabilitas soal tersebut, dari kedua hal ini dapat digunakan untuk memberikan penilaian terhadap kemampuan responden.

Dalam penelitian ini akan dibahas mengenai analisis dan evaluasi terhadap alat tes yang dikonstruksi agar memiliki daya ukur yang tinggi dan tidak menimbulkan miskonsepsi jika dikerjakan remaja yang memiliki latar belakang keilmuan yang berbeda-beda. Alat ukur ini dikembangkan dengan harapan bersifat generik atau dapat diberlakukan di semua kalangan remaja dan dapat mengases keterampilan berpikir kritis dengan baik.

METODE

Pendekatan yang digunakan adalah pendekatan kuantitatif dengan metode pengembangan instrumen tes dengan analisis deskriptif terhadap hasil uji coba tes keterampilan berpikir kritis remaja (TKBKR). Pengumpulan data dilakukan dengan tes berupa TKBKR yang dikonstruksi oleh peneliti yang memuat 36 soal dari 6 bacaan. Subyek penelitian ini adalah remaja berusia 19-22 tahun pada program studi Pendidikan Biologi tingkat pertama di Universitas Pendidikan Indonesia.

Teknik analisis data menggunakan analisis pemodelan RASCH. Penelitian ini mengidentifikasi beberapa hal yakni meliputi; 1) reliabilitas, 2) validitas, 3) analisis tingkat kesulitan soal, 4) sebaran item, serta 5) gambaran proyeksi jawaban responden dalam skalogram. Kelima hal tersebut cukup menggambarkan daya ukur instrumen terhadap keterampilan berpikir kritis mahasiswa. Selain itu, analisis ini akan menunjukkan abilitas responden dalam mengerjakan tes yang nantinya akan digunakan sebagai data untuk mengukur keterampilan berpikir kritis mahasiswa pendidikan biologi.

HASIL DAN PEMBAHASAN

Sebelumnya saya bahas alasan mengapa menggunakan analisis RASCH sebagai analisis data yang saya gunakan saat ini. Berdasarkan teori yakni yang menyatakan bahwa analisis RASCH dapat digunakan untuk sampel sebesar 30 hingga 300 responden (Bon and Fox, 2007), sedangkan IRT (*Item Response Theory*) digunakan untuk sampel sebesar 200 hingga 1000 (Seon, 2009; Istiyono, 2014). Berdasarkan teori tersebut, jika dalam penelitian ini saya hanya menggunakan responden sebesar 29 mahasiswa atau mendekati 30 orang, maka saya lebih sesuai menggunakan analisis pemodelan RASCH untuk mengevaluasi instrumen TKBKR ini.

Analisis RASCH memiliki beberapa kelebihan yaitu meliputi lima prinsip model pengukuran yakni; 1) mampu memberikan skala linier dengan interval yang sama; 2) dapat melakukan prediksi terhadap data yang tidak terbaca; 3) mampu memberikan estimasi yang lebih tepat; 4) dapat mendeteksi ketidaktepatan model; dan 5) dapat menghasilkan pengukuran yang *replicable* (Sumintono & Widhiarso, 2013).

Hasil penelitian ini menggambarkan beberapa komponen yang dapat diidentifikasi meliputi reliabilitas, daya ukur instrumen terhadap keterampilan berpikir kritis mahasiswa, validitas instrumen, tingkat kesulitan soal, serta kualitas instrumen dalam mengukur beberapa kelompok subyek. Penjelasan terperinci dapat diperhatikan beberapa hasil analisis pemodelan RASCH berikut ini.

SUMMARY OF 29 MEASURED PERSON

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	41.1	36.0	-.96	.22	1.00	-.1	1.01	-.1
S.D.	7.8	.0	.39	.01	.34	1.6	.41	1.4
MAX.	57.0	36.0	-.14	.25	1.87	3.6	2.25	3.7
MIN.	16.0	36.0	-2.25	.22	.34	-3.9	.37	-3.3
REAL RMSE	.24	TRUE SD	.31	SEPARATION	1.31	PERSON RELIABILITY	.63	
MODEL RMSE	.22	TRUE SD	.32	SEPARATION	1.43	PERSON RELIABILITY	.67	
S.E. OF PERSON MEAN = .07								

PERSON RAW SCORE-TO-MEASURE CORRELATION = 1.00

CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .65

SUMMARY OF 35 MEASURED ITEM

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	34.1	29.0	.00	.25	1.00	.0	1.01	.0
S.D.	14.2	.0	.79	.04	.27	1.2	.30	1.1
MAX.	54.0	29.0	1.79	.37	1.50	2.3	1.73	2.0
MIN.	6.0	29.0	-1.20	.22	.50	-2.2	.50	-2.3
REAL RMSE	.27	TRUE SD	.75	SEPARATION	2.82	ITEM RELIABILITY	.89	
MODEL RMSE	.25	TRUE SD	.75	SEPARATION	2.98	ITEM RELIABILITY	.90	
S.E. OF ITEM MEAN = .14								

MAXIMUM EXTREME SCORE: 1 ITEM
 UMEAN=.0000 USCALE=1.0000

SUMMARY OF 36 MEASURED (EXTREME AND NON-EXTREME) ITEM

Gambar 2. Uji Reliabilitas

Pada gambar di atas dapat dijelaskan bahwa reliabilitas dalam analisis pemodelan RASCH dibagi menjadi tiga bagian, yaitu 1) reliabilitas keseluruhan, 2) reliabilitas item, dan 3) reliabilitas responden. Dapat dibaca pada bagian analisis Alpha Cronbach tertera koefisien reliabilitas sebesar 0,65, sedangkan pada *person reliability* sebesar 0,63 dan pada *item reliability* sebesar 0,89 atau dalam katori tinggi.

Pada gambar di atas menjelaskan bahwa nilai *mean measure* pada person sebesar -0,96 . Nilai tersebut menunjukkan bahwa kecenderungan abilitas subyek lebih kecil dibandingkan dengan tingkat kesulitan soal. Lalu pada nilai *Alpha Cronbach* menunjukkan koefisien reliabilitas yang dihitung berdasarkan teori tes

klasik dengan hasil koefisien sebesar 0,65. Hal ini menunjukkan reliabilitas TKBKR ini secara umum belum memuaskan dalam uji coba skala kecil.

Nilai *person reliability* sebesar 0,63, sedangkan nilai *item reliability* sebesar 0,89. Hal ini menunjukkan bahwa subyek memberikan jawaban belum konsisten, namun kualitas butir soal pada TKBKR memiliki reliabilitas yang cukup baik. Nilai *separation* dalam gambar di atas juga dapat diinterpretasikan sebagai kualitas instrumen dan juga kualitas subyek penelitian kita. Pada gambar di atas jika digunakan untuk menghitung kualitas instrumen dapat dilihat nilai *separation* pada *item separation* dengan hitungan $H = [(4 \times 2,82) + 1] / 3$ yakni 4,09 atau dibulatkan

menjadi 4, yang berarti terdapat 4 kelompok butir soal. Lalu pada nilai *person separation* terdapat nilai 1,31. Jika dihitung dengan rumus $H = [(4 \times 1,31) + 1] / 3$ yakni 2,08 atau dibulatkan menjadi 2, sehingga dapat diartikan bahwa instrumen ini hanya dapat mengukur 2 kelompok saja. Jadi, pada instrumen ini dapat digunakan untuk mengukur subyek yang memiliki keterampilan berpikir kritis pada kategori sedang dan rendah saja pada subyek uji coba pertama ini.

Tes ini memiliki reliabilitas kurang memuaskan diakibatkan karena performa *testee*

saat mengerjakan tes sangat lemah karena di dalam kondisi setelah kuliah pada jam ketiga menjelang siang, sehingga banyak keluhan lapar dan pengen istirahat. Selain itu, bisa jadi karena akibat responden dalam memberikan jawaban tidak konsisten. Hal ini disebut dengan faktor *events producing inconsistent performance* (Friedenberg, 1995). Hal ini dapat dibuktikan pada hasil analisis RASCH pada gambar berikut.

GUTTMAN SCALOGRAM OF RESPONSES:

PERSON	ITEM	
		1112 233313 2 22312 11 223322 111
		212687712604143923816755360550489349

17	+122222222222121220220222122200122220	00017P
10	+22201022222222220022302022222110000	00010P
18	+22322202222230322221210000102010210	00018P
23	+320220332222200231222230002110200020	00023P
14	+32322332202222202202011200100202000	00014P
21	+23212311222221222221222020010010000	00021P
25	+12222232122030221221202212021000000	00025P
26	+12221211022212222221220022310100000	00026P
2	+20222222220222120220221022011100010	00002P
4	+22211222221220221220212011000020200	00004P
11	+22222222101202201102022320121100000	00011P
12	+0122222221222220021202201002101000	00012P
22	+222120201122232221203000200021020000	00022P
3	+22222212211223122002000210011000000	00003P
8	+22112221222202122003200002011010000	00008P
24	+220221232222020110111022210022000000	00024P
1	+222210022221203122021201001001012000	00001P
9	+122222210222021221101221012001000000	00009P
15	+12222222222131001102110000300200000	00015P
5	+22222212211212122020100001100020000	00005P
20	+32222222121202021022101200000000000	00020P
7	+20222222122020202001202020010101000	00007P
13	+22222220211121202000020200212000000	00013P
6	+22222212121212111210010011100000000	00006P
19	+32222212221220220211100110000000000	00019P
16	+222223212202020021201012010000000000	00016P
28	+221202010212010001120203200200203000	00028P
27	+222210222000221002120010100000010100	00027P
29	+11111011010111001100100000010000000	00029P

		1112 233313 2 22312 11 223322 111

Gambar 3. Skalogram Mahasiswa Pendidikan Biologi

Pada gambar 3, diketahui pola respon subyek terhadap tes. Skor jawaban dominan pada angka 2, sedangkan angka 0 dan angka 1 memiliki frekuensi yang tidak beraturan sehingga memengaruhi reliabilitas. Selain itu, adanya kemungkinan responden memiliki miskonsepsi

terhadap soal karena kemampuan pemahaman yang kurang dan penyelesaian soal dengan kondisi kurang konsentrasi. Beberapa soal yang tidak valid juga mempengaruhi reliabilitas seperti halnya gambar berikut.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	TOTAL MEASURE	MODEL		INFIT		OUTFIT		PT-MEASURE		EXACT OBS%	MATCH EXP%	ITEM	
				S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.					
9	0	29	4.47	1.80	MAXIMUM MEASURE										Q9
14	6	29	1.79	.37	1.00	.2	.64	-.6	.51	.16	86.2	83.9	Q14		
13	7	29	1.66	.35	1.20	.6	1.02	.2	.33	.18	86.2	80.2	Q13		
19	11	29	1.27	.28	1.48	1.4	1.73	1.7	.12	.22	62.1	63.5	Q19		
8	14	29	1.05	.26	.93	-.2	.94	-.1	.27	.24	51.7	52.7	Q8		
4	15	29	.98	.25	.87	-.5	.90	-.3	.32	.24	69.0	51.8	Q4		
20	18	29	.80	.24	.93	-.3	.88	-.4	.22	.26	58.6	39.2	Q20		
25	19	29	.75	.23	.90	-.5	.96	-.1	.15	.27	44.8	32.8	Q25		
35	20	29	.69	.23	1.28	1.4	1.27	1.2	.28	.27	31.0	29.7	Q35		
26	21	29	.64	.23	.85	-.8	.81	-.9	.42	.27	31.0	29.7	Q26		
30	21	29	.64	.23	.91	-.4	.86	-.6	.47	.27	31.0	29.7	Q30		
23	26	29	.39	.22	1.39	2.3	1.35	1.8	.12	.29	10.3	27.0	Q23		
5	27	29	.35	.22	1.22	1.4	1.22	1.2	.17	.29	27.6	26.8	Q5		
15	28	29	.30	.22	.94	-.3	.91	-.4	.47	.29	31.0	26.9	Q15		
6	32	29	.11	.22	1.16	1.0	1.17	1.0	.27	.30	41.4	26.8	Q6		
17	32	29	.11	.22	1.04	.3	1.04	.3	.31	.30	13.8	26.8	Q17		
21	33	29	.06	.22	1.02	.2	1.00	.1	.40	.30	17.2	26.6	Q21		
18	34	29	.02	.22	.77	-1.4	.76	-1.4	.44	.30	27.6	26.5	Q18		
33	36	29	-.08	.22	.96	-.2	1.04	.3	-.20	.31	44.8	33.2	Q33		
22	37	29	-.13	.22	1.14	.8	1.15	.8	.43	.31	24.1	34.7	Q22		
29	39	29	-.23	.22	.62	-2.2	.58	-2.3	.67	.31	51.7	37.6	Q29		
3	41	29	-.33	.23	1.40	1.7	1.45	1.8	.15	.30	41.4	40.5	Q3		
1	42	29	-.38	.23	1.08	.4	1.09	.4	.24	.30	37.9	40.8	Q1		
24	42	29	-.38	.23	1.50	2.0	1.55	2.0	.09	.30	37.9	40.8	Q24		
34	44	29	-.50	.24	.57	-2.0	.56	-1.9	.51	.30	58.6	46.8	Q34		
10	45	29	-.55	.24	.61	-1.7	.57	-1.8	.49	.30	72.4	51.4	Q10		
36	46	29	-.61	.25	.64	-1.5	.66	-1.2	.41	.30	62.1	54.9	Q36		
32	48	29	-.74	.26	.91	-.2	.87	-.3	.25	.29	75.9	62.2	Q32		
7	49	29	-.81	.26	1.50	1.5	1.65	1.8	.14	.29	69.0	66.0	Q7		
27	49	29	-.81	.26	1.09	.4	1.19	.7	.33	.29	65.5	66.0	Q27		
31	49	29	-.81	.26	.88	-.3	.93	-.1	.33	.29	51.7	66.0	Q31		
28	50	29	-.88	.27	.50	-1.8	.50	-1.7	.30	.29	82.8	67.1	Q28		
16	51	29	-.95	.27	.62	-1.2	.74	-.7	-.02	.28	79.3	68.1	Q16		
11	53	29	-1.11	.29	.80	-.5	.87	-.2	.22	.27	86.2	71.2	Q11		
12	53	29	-1.11	.29	.95	.0	1.10	.4	.24	.27	79.3	71.2	Q12		
2	54	29	-1.20	.30	1.21	.7	1.37	1.0	.03	.27	65.5	72.0	Q2		

Gambar 4. Analisis Item Measure

Pada gambar di atas diketahui ada satu soal yang berada pada "maximum measure" artinya soal nomor 9 tidak dapat dipertahankan karena terlalu sulit dan tidak ada satupun mahasiswa yang dapat menjawab soal tersebut. Soal yang masih dapat dipertahankan untuk diujikan kembali tinggal 35 nomor, akan tetapi setelah diuji validitas berdasarkan norma soal yang memiliki daya beda $0,4 < Pt \text{ Measure Corr} < 0,85$ menunjukkan bahwa soal tersebut dapat dipertahankan. Jika ingin mengetahui soal yang dapat dipahami dengan baik oleh mahasiswa dan tidak menimbulkan miskonsepsi dalam jawaban perlu diperhatikan juga kolom *Outfit MNSQ* ($0,5 < MNSQ < 1,5$) dan *Outfit ZSTD* ($-2,0 < ZSTD > 2,0$) beserta norma pada *Pt. Measure Corr* (Sumintono & Widhiarso, 2015). Jadi berdasarkan gambar hasil analisis di atas, soal yang perlu dihilangkan karena tidak valid dan tidak memiliki daya beda yang baik yaitu nomor

09, 24, 19, 08, 20, 25, 35, 23, 05, 06, 33, 03, 01, 32, 07, 16, 11, 12, dan 02. Berdasarkan item yang harus dibuang berjumlah 19 nomor, selebihnya masih dapat direvisi untuk perbaikan sebelum diterbitkan.

SIMPULAN

Instrumen Tes Keterampilan Berpikir Kritis Remaja (TKBKR) jika diujikan ke mahasiswa Pendidikan Biologi Universitas Pendidikan Indonesia pada angkatan pertama menghasilkan reliabilitas sedang dengan koefisien 0.65. Reliabilitas item memiliki kategori tinggi dan reliabilitas responden dalam menjawab berada pada kategori sedang.

Kecenderungan abilitas responden dalam mengerjakan soal lebih kecil jika dibandingkan dengan tingkat kesulitan soal. Nilai reliabilitas tes secara keseluruhan belum memuaskan karena

koefisien reliabilitas masih di bawah 0.80. Kualitas butir soal yang dikonstruksi cukup baik, sedangkan konsistensi jawaban dari subyek masih belum memuaskan. *Item separation* menunjukkan kualitas instrumen masih kurang baik, karena hanya terdapat 4 kelompok butir soal yang dapat digunakan untuk mengidentifikasi 2 kelompok subyek saja.

UCAPAN TERIMA KASIH

Ucapan terima kasih ini ditujukan untuk kedua pembimbing saya, yaitu Dr. Ilfiandra, M.Pd. dan Dr. Nurhuda, M.Pd. yang telah memberikan arahan dan bimbingan kepada saya hingga selesainya pelaporan penelitian ini. Serta ketiga *expert* yang sudah memvalidasi instrumen ini hingga terselesaikan penelitian tepat waktu. Berikutnya, ucapan terima kasih ini saya tujukan kepada mahasiswa saya prodi Pendidikan Biologi Universitas Pendidikan Indonesia angkatan 2017 yang telah berpartisipasi dalam penelitian ini.

REFERENSI

- Azwar, S. (2016). *Konstruksi Tes Kemampuan Kognitif*. Yogyakarta : Pustaka Pelajar.
- Bensley, D. A., & Murtagh, M. P. (2012a). Guidelines for a Scientific Approach to Critical Thinking Assessment. *Society for The Teaching of Psychology*, 39(1), 5–16. <https://doi.org/10.1177/0098628311430642>
- Bensley, D. A., & Murtagh, M. P. (2012b). Teaching of Psychology. *Sage Publisher*, 39(1), 5–16. <https://doi.org/10.1177/0098628311430642>
- Bond, T. G., & Fox, C.M. (2015). *Applying The Rasch Model Fundamental Measurement in The Human Sciences, Third Edition*. New York : Routledge.
- Bond, TG and Fox, CM. (2007). *Applying the Rasch Model. Fundamental Measurement in the Human Sciences (2 rd edition)*. Mahwah, NJ: Lawrence Erlbaum. Bonk.
- Celuch, K. (2009). Student Self-Identity as a Critical Thinker The Influence of Attitudes , Attitude Strength , and Normative Beliefs, 31–39.
- Facione, P. A. (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (Research Findings and Recommendations)*. Newark, De: American Philosophical Association.
- Fisher, Alec. (2009). *Berpikir Kritis (Sebuah Pengantar)*. Alih Bahasa: Benyamin Hadinata. Jakarta : Erlangga.
- Freidenberg, L. (1995). *Psychological Testing: Design, Analysis, and Use*. Massachusetts : Allyn & Bacon.
- Garrett, M. L. (2013). An Examination of Critical Thinking Skills in High School Choral Rehearsals. <https://doi.org/10.1177/0022429413497219>
- Heinrich, W. F., Habron, G. B., Johnson, H. L., & Goralnik, L. (2015). Critical Thinking Assessment Across Four Sustainability-Related Experiential Learning Settings. <https://doi.org/10.1177/1053825915592890>
- Istiyono, E. (2014). Developing Higher Order Thinking Skill Test of Physics (PhysTHOTS) For Senior High School Students. *Penelitian Dan Evaluasi Pendidikan*, 18(1), 1–12. Retrieved from <http://journal.uny.ac.id/index.php/jpep/article/view/2120>
- Kargar, F. R., Ajilchi, B., Goreyshi, M. K., & Noohi, S. (2013). Effect of Creative and Critical Thinking Skills Teaching on Identity Styles and General Health in Adolescents. *Procedia - Social and Behavioral Sciences*, 84(2003), 464–469. <https://doi.org/10.1016/j.sbspro.2013.06.585>
- Marfu'i, Lucky NR. (2018). *Pengembangan Tes Keterampilan Berpikir Kritis pada Mahasiswa Universitas Pendidikan Indonesia*. Thesis Universitas Pendidikan Indonesia, Bandung.
- Marfu'i, Lucky NR. (2018). The Analysis of Critical Thinking Skills Test in Social-Problems for Physics Education Students with Rasch Model. *MSCEIS Conference*. Bandung, West Java- Indonesia.
- Miri, Barak, Ben- Chaim, And Uri, Zoller. (2007). Purposely Teaching for the Promotion of Higer-Order Thinking Skills: A Case of Critical Thinking. *Research Science Education Journal- Springer Science* 37, Pp. 353-369.
- Seon, Hi Sin. (2009). How to tread omitted respons in Rasch model based equa- ting [Versi elektronik] Practical Assess- ment, Research & Evaluation. ISSN 1531- 7714, Volume 14, Number 1, p: 1-8
- Suarsana, I. M., And G.A. Mahayukti. (2013). Pengembangan E-Modul Berorientasi Pemecahan Masalah untuk Meningkatkan Keterampilan Berpikir Kritis Mahasiswa. *Jurnal Pendidikan Indonesia*, Vol. 2, No.2, 264-275.
- Sumintono, B & Widhiarso, W. (2015). *Aplikasi pemodelan RASCH pada assessment pendidikan*. Cimahi : Trim Komunikata.
- Sumintono, B & Widhiarso, W. (2013). *Aplikasi Model Rasch untuk Penelitian Ilmu-Ilmu Sosial*. Cimahi : Trimkom Publishing House/