



PENGEMBANGAN INSTRUMEN TES KETERAMPILAN BERPIKIR KRITIS DALAM BENTUK PILIHAN GANDA BERALASAN (POLITOMUS) DI DIY

F.S. Putri✉, E. Istiyono, E. Nurcahyanto

Pascasarjana Program Studi Pendidikan Fisika, Universitas Yogyakarta
Indonesia, 55281

Info Artikel

Sejarah Artikel:

Diterima April 2016
Disetujui April 2016
Dipublikasikan Agustus
2016

Keywords:

*Development of instrument,
critical thinking skills test,
polytomus, PCM*

Abstrak

Penelitian ini adalah penelitian pengembangan instrumen tes keterampilan berpikir kritis dalam bentuk pilihan ganda beralasan (politomus). Keterampilan berpikir kritis memuat empat aspek yaitu menginterpretasi, menganalisis, menginferensi, dan menjelaskan yang selanjutnya diturunkan menjadi 28 indikator. Instrumen terdiri dari 50 item yang terbagi dalam 2 perangkat tes. Validasi instrumen dilakukan oleh 2 ahli pengukuran, 1 ahli pendidikan fisika, dan 1 praktisi. Instrumen telah diujicobakan pada 262 peserta didik yang tersebar di 5 SMA di Daerah Istimewa Yogyakarta. Data politomus empat kategori dari hasil pengukuran dianalisis mengikuti *Partial Credit Model* (PCM) menggunakan program Quest. Seluruh item dinyatakan fit dengan model berdasarkan kriteria batas terendah dan tertinggi INFIT MNSQ 0,77 sampai dengan 1,30. Reliabilitas estimasi item sebesar 0,85 yang tergolong kategori tinggi. Tingkat kesulitan item berada pada rentang 0,00 sampai 0,24 yang berarti item dalam kategori baik karena berada di antara -2,0 sampai 2,0.

Abstract

This research is categorized as an research and development which focused on critical thinking test instrument in reasoned multiple choice test (polytomus). Critical thinking skills covers four issues; interpreting, analyzing, inferencing, and explanating which continuing in 28 indicators. The instruments consists of 50 items which alloted in two set of tests. The validation of the instrument conducted by measurement experts in pairs, an physics educational expert, and one sholars. Try-out of instrument already tested on 262 students around the five senior high school in Yogyakarta. The fourth of politomus data categorized based on analysis measurement-Partial Credit Model (PCM) and use Quest Programs. All of th items stated on fit with model based on criteria low boundaries and the highest INFIT MNSQ 0,77 until 1.30. The reliability item on point 0.85 which categorized in high boundaries. The difficulty level place on more or less 0.00 – 0.24, it means the items in good categorized because on -2.0 until 2.0.

PENDAHULUAN

Perkembangan pendidikan di Indonesia mengikuti acuan tolak ukur pendidikan di tingkat internasional. *Partnership for 21st Century Skills* (2013) sebagai salah satu acuan pendidikan menyebutkan bahwa kompetensi yang perlu ditingkatkan pada peserta didik di abad-21 meliputi materi inti, keterampilan belajar dan berinovasi, keterampilan teknologi informasi dan media, dan keterampilan hidup dan karier (*survival skill*). Wagner (2008) menekankan delapan *survival skills* yang memiliki nilai penting di era abad ke-21 yaitu: (1) *communication skills*; (2) *critical and creative thinking*; (3) *inquiry/reasoning skills*; (4) *interpersonal skills*; (5) *multicultural/multilingual literacy*; (6) *problem solving*; (7) *information/digital literacy*; dan (8) *technological skills*. Keterampilan-keterampilan tersebut sepatutnya dimiliki oleh peserta didik di Indonesia agar dapat bersaing di tingkat internasional.

Permendikbud No.64 Tahun 2013 tentang standar isi mengkerucutkan pada pengembangan salah satu *survival skill* tersebut. Pada peraturan tersebut dinyatakan bahwa salah satu kompetensi yang harus dikembangkan dalam pelaksanaan kurikulum 2013 adalah kompetensi berpikir kritis. Namun, keterampilan berpikir kritis peserta didik di Indonesia belum dikembangkan dengan maksimal. Hal tersebut dalam dilihat pada hasil PISA (*Programme of International Student Assessment*) yang salah satu komponennya adalah keterampilan berpikir kritis. OECD (*Organization for Economic Cooperation and Development*) menyatakan bahwa Indonesia menduduki peringkat 69 dari 76 negara pada PISA. Hasil ini mengindikasikan bahwa Indonesia masih menduduki peringkat rendah di dunia. Oleh karena itu, pengembangan keterampilan berpikir kritis adalah tantangan yang harus dihadapi pendidikan untuk memenuhi tuntutan abad ke-21 dan peraturan pemerintah tersebut.

Ennis (1996: 50) memberikan definisi berpikir kritis adalah berpikir reflektif yang berfokus pada pola menyajikan keputusan tentang apa yang diyakini dan harus dilakukan. Johnson (2002: 86) juga menyatakan bahwa berpikir kritis merupakan sebuah proses berpikir untuk membangun perkembangan kepercayaan dan mengambil tindakan dengan cara terorganisasi dalam kegiatan mental seperti memecahkan masalah, mengambil keputusan, menyisati, menganalisis asumsi, dan melakukan penelitian. Dengan demikian, seseorang dikatakan berpikir kritis dapat tercermin dari cara berpikir yang terarah dan terorganisir dengan baik dalam menyelesaikan suatu permasalahan.

Pembelajaran fisika sebagai bagian dari sains mengacu pada pembelajaran yang berorientasi pada proses penemuan. Salah satu cara untuk menemukan konsep adalah melakukan percobaan yang di dalamnya menggunakan metode saintifik. Dewey, *et al.* (1998) menyatakan bahwa keterampilan berpikir kritis dapat diasosiasikan dengan scientific thinking yang diindikasikan dengan lima indikator. Kelima indikator tersebut yaitu keterampilan dalam identifikasi masalah, keterampilan dalam formulasi hipotesis, keterampilan dalam pengumpulan data, proses analisis dan evaluasi, dan keterampilan dalam menyusun kesimpulan. Kelima indikator tersebut dilakukan peserta didik dengan berurutan atau sistematis.

Selain itu, penelitian yang dilakukan oleh Miri, *et al.* (2007: 365) telah mengembangkan beberapa indikator keterampilan berpikir kritis yaitu : (1) mengidentifikasi masalah; (2) mendesain eksperimen; (3) mengumpulkan perlengkapan yang dimiliki; (4) menyusun hipotesis; (5) memspesifikasikan instrumen yang akan digunakan. Penelitian mendapatkan hasil bahwa terdapat perbedaan keterampilan berpikir kritis peserta didik di kelas kontrol dan kelas eksperimen.

Biasanya, proses pengukuran keterampilan berpikir kritis dilakukan dengan objektif dalam bentuk pilihan ganda. Namun, berdasarkan penelitian Ku (2009), soal pilihan ganda tidak dapat mengkuantitatifkan keterampilan berpikir kritis. Oleh karena itu keterampilan berpikir kritis yang dikembangkan dalam penelitian ini dalam bentuk instrumen pilihan ganda yang disusun dengan format multi respon (dengan disertai alasan). Alasan juga dibuat dalam bentuk pilihan ganda. Hal ini bertujuan agar peserta didik memberikan alasan dalam pemilihan jawaban.

Berdasarkan penelitian yang telah dilakukan sebelumnya tersebut, maka instrumen tes yang disusun dalam penelitian ini harus memuat indikator-indikator yang operasional

METODE PENELITIAN

Penelitian adalah penelitian pengembangan instrumen tes yang dikembangkan oleh Oriondo & Antonio (1984: 34). Pengembangan instrument tes mengikuti sistematika pengembangan dengan tahap: (1) perencanaan; (2) uji coba; (3) validasi instrumen; (4) pengukuran reliabilitas; dan (5) proses interpretasi skor. Pada tahap perencanaan tes terdapat lima langkah yang dilakukan yaitu (1) penentuan tujuan tes; (2) pengembangan bentuk tes; (3) pembuatan kisi-kisi tes; (4) penulisan tes; dan (5) revisi tes.

Tes yang dikembangkan bertujuan untuk mengukur keterampilan berpikir kritis peserta didik dalam bentuk pilihan ganda berbasalan. Pengembangan tes dibatasi pada mata pelajaran fisika materi hukum Newton. Tes dikembangkan berdasarkan empat aspek keterampilan berpikir kritis dan selanjutnya dijabarkan dalam tujuh indikator. Indikator yang telah disusun kemudian dikembangkan ke dalam butir soal sebanyak 50 item.

Pada pengembangan instrumen test ini dibantu oleh 2 orang pakar pengukuran, 1 pakar pendidikan, dan 1 orang praktisi. Instrumen yang telah melalui ijustifikasi pakar selanjutnya diuji kualitas dari instrumen yang telah disusun.

untuk mengukur indikator ketercapaian keterampilan berpikir kritis peserta didik. Keterampilan berpikir kritis yang dikembangkan memuat lima aspek, yaitu: (1) menginterpretasi; (2) menganalisis; (3) menginferensi; (4) menjelaskan. Keempat aspek tersebut diturunkan menjadi tujuh sub aspek sebagai berikut : (1) mengidentifikasi masalah; (2) mengintepretasi data dalam percobaan; (3) membuat solusi permasalahan; (4) menyusun hipotesis; (5) mendesain percobaan; (6) membuat kesimpulan; (7) mengidentifikasi kesalahan dalam percobaan. Selanjutnya setiap sub aspek tersebut diturunkan menjadi 4 indikator.

Pada tahap ini subjek uji coba instrument melibatkan 262 peserta didik tingkat sekolah menengah di Daerah Istimewa Yogyakarta yang terdiri dari 6 SMA. Hasil dari uji lapangan ini selanjutnya di analisis dengan menggunakan teori respon butir dengan pendekatan model Rasch yang dianalisis dengan menggunakan Software Quest.

Analisis data dilakukan dengan menggunakan program Quest untuk menguji kecocokan model, tingkat kesulitan item, dan reliabilitas item. Menurut Adam & Khoo (1996: 30), item lolos dalam pengujian *goodness of fit* berdasarkan nilai rerata INFIT *Mean of Square* (Mean INFITMNSQ) dan simpangan bakunya. Jika rerata INFIT MNSQ sekitar 1,0 dan simpangan bakunya 0,0 atau rerata INFIT t mendekati 0,0 dan simpangan bakunya 1,0, maka keseluruhan tes fit dengan model. Suatu item atau *testi/case/person* dinyatakan fit dengan model dengan batas kisaran INFIT MNSQ dari 0,77 sampai 1,30. Selain itu, item dikatakan baik jika indeks kesukaran lebih dari -2.0 atau kurang dari 2.0 (Hambleton & Swaminathan, 1985: 36).

HASIL DAN PEMBAHASAN

Analisis Kualitas Instrumen Tes Keterampilan Berpikir Kritis

Instrumen tes dikembangkan berdasarkan analisis aspek, sub aspek, dan indikator. Setiap sub aspek diturunkan menjadi empat indikator yang selanjutnya diturunkan menjadi butir/item soal. Dalam menanggulangi adanya item yang gugur, terdapat beberapa indikator yang disusun menjadi 2 butir. Distribusi jumlah item berdasarkan aspek dan sub aspek dapat dilihat pada tabel 1. Soal dibagi menjadi 2 paket yang di dalamnya terdapat 6 item *anchor*. Setiap paket terdiri dari 28 butir, sehingga item keseluruhan yang dikembangkan berjumlah 50 item.

Tabel 1. Distribusi Jumlah Item Berdasarkan Aspek dan Sub Aspek

No	Aspek	Sub Aspek	Jumlah Item
1.	Mengintepretasi	mengidentifikasi masalah	7
		mengintepretasi data dalam percobaan	6
2.	Menganalisis	membuat solusi permasalahan	8
3.	Menginferensi	merumuskan hipotesis	8
		mendesain percobaan	8
		membuat kesimpulan	8
4.	Menjelaskan	mengidentifikasi kesalahan dalam percobaan	5

Seluruh item tersebut ditelaah oleh beberapa ahli. Pada analisa tiap item keterampilan berpikir kritis berdasarkan ahli, didapatkan 8 item harus direvisi dan selebihnya dinyatakan valid. Setelah item direvisi, seluruh item divalidasi dan mendapatkan hasil dari kesepakatan interater menggunakan persamaan Aiken's V sebesar 0.97. Hasil tersebut

merupakan hasil rata-rata dari keseluruhan item maka setiap item dari instrumen dinyatakan layak digunakan.

Pada tahap validasi instrumen dan analisis kualitas butir instrumen dengan menggunakan teori respon butir. Berdasarkan pendekatan model rasch didapatkan 50 butir instrumen fit dengan model karena memenuhi kriteria infit mnsq sebesar 0.77 sampai dengan 1.30 seperti ditunjukkan pada gambar 1. Hal ini dapat disimpulkan bahwa setiap item instrumen telah fit digunakan untuk mengukur keterampilan berpikir kritis secara empirik.

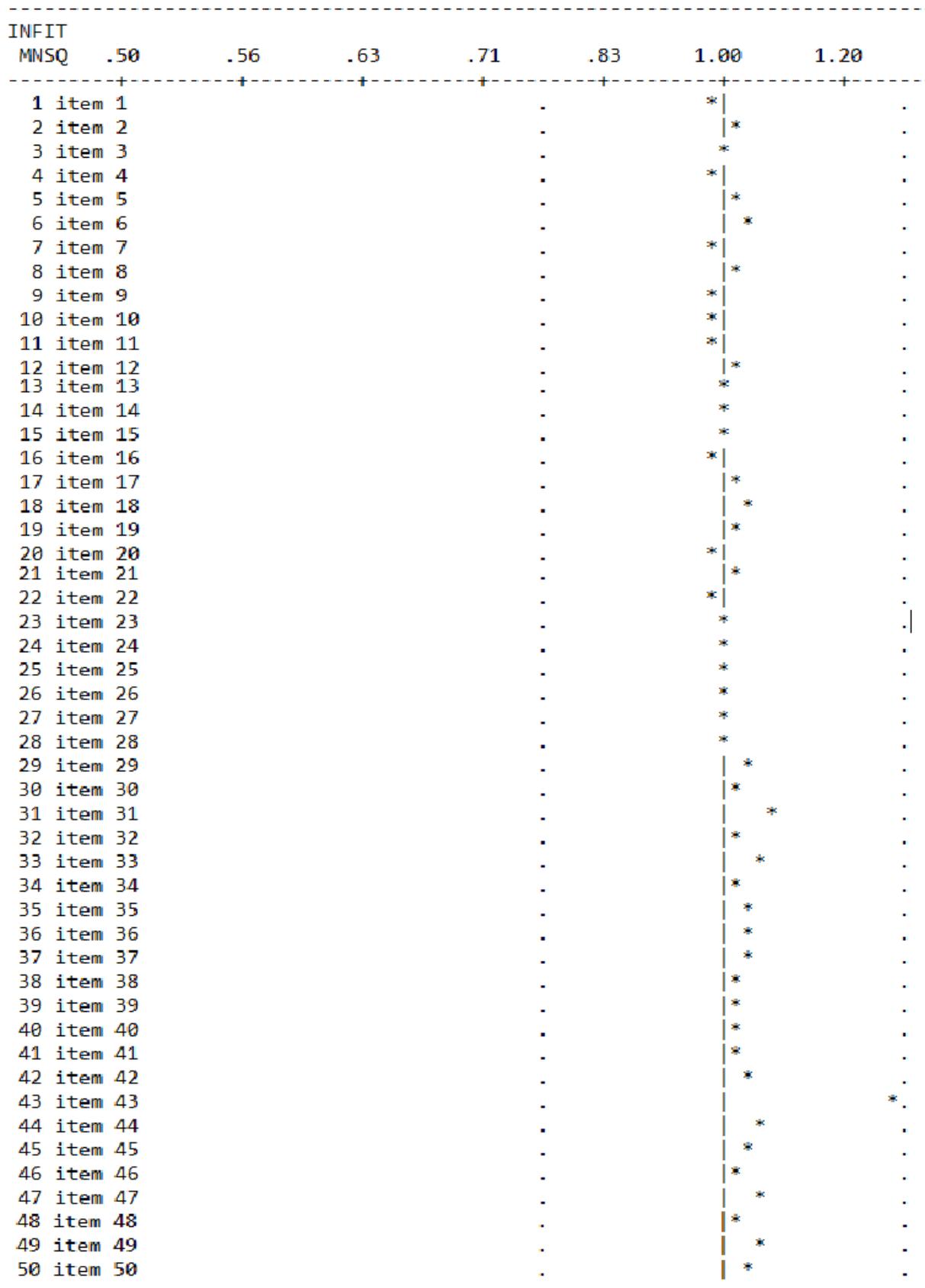
Tabel 2. Hasil Uji Validitas Empirik Instrumen Tes Keterampilan Berpikir Kritis

No.	Parameter	Estimasi Butir	Estimasi Case
1.	INFIT MNSQ	1.02 ± 0.04	1.02 ± 0.14
2.	OUTFIT MNSQ	1.02 ± 0.05	1.02 ± 0.14
3.	Difficulty rata-rata	0.0 ± 0.24	
4.	Reability Estimasi	0.85	

Tabel 2 juga menunjukkan bahwa tingkat kesulitan yang dapat diukur oleh instrumen ada pada jangkuan -2 sampai dengan +2 yaitu 0.00 ± 0.24 sehingga instrumen baik untuk digunakan mengukur keterampilan berpikir kritis.

Hasil ujicoba instrumen secara singkat disajikan dalam gambar 1. Instrumen fit digunakan untuk mengukur keterampilan berpikir kritis dengan nilai INFIT *meansquare* berada pada 1.02 ± 0.04 .

Uji Reliabilitas item secara klasik dilakukan dengan perhitungan alpha conbrach dengan menggunakan program Quest. Berdasarkan analisis diperoleh nilai reliabilitas sebesar 0.85, hal ini menunjukkan bahwa tingkat reliabilitas instrumen berada dalam kategori tinggi.



Gambar 1. INFIT Meansquare Instrumen Keterampilan Berpikir Kritis

Analisis Keterampilan Berpikir Kritis Peserta didik

Tes keterampilan berpikir kritis dalam penelitian ini dikembangkan dengan tipe pilihan ganda beralasan. Kriteria penskoran tes dibuat dengan empat kategori. Kategori 1 jika peserta didik menjawab soal salah dan jawaban salah. Kategori 2 jika peserta didik menjawab soal benar dan alasan salah. Kategori 3 jika peserta didik menjawab soal salah dan alasan benar. Kategori 4 jika peserta didik menjawab soal

benar dan alasan benar. Oleh karena itu dapat diketahui tingkat berpikir kritis peserta didik dari kategori tersebut.

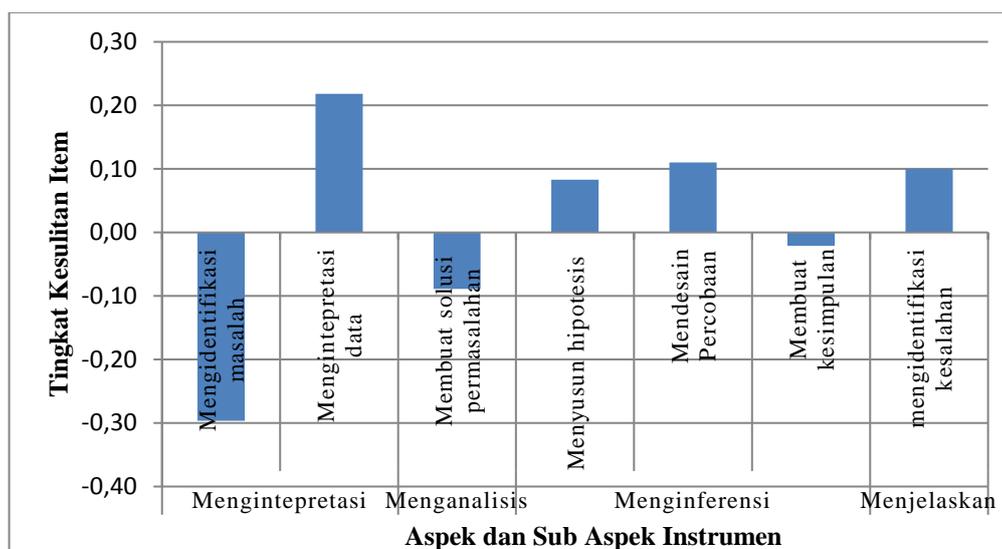
Berdasarkan hasil analisis tingkat kesulitan butir diketahui bahwa item tes keterampilan berpikir kritis memiliki perbedaan pada masing-masing aspek dan sub aspek. Secara ringkas tingkat kesulitan butir untuk setiap aspek dan sub aspek dapat dilihat dari bagan Tabel 3.

Tabel 3. Tingkat Kesulitan Butir Masing-masing Sub Aspek Kegiatan Uji coba

Aspek	Sub Aspek	Difficulty	Kesulitan Tahap			
			Kategori 1	Kategori 2	Kategori 3	Kategori 4
Menganalisis	Mengidentifikasi masalah	-0,30	0,35	-0,46	0,56	-0,45
	Menginterpretasi data	0,22	1,05	-0,56	-0,05	-0,88
Menginferensi	Membuat solusi permasalahan	-0,09	0,93	-0,56	0,34	-0,44
	Menyusun hipotesis	0,08	0,46	-0,25	0,55	-0,88
Menjelaskan	Mendesain Percobaan	0,11	0,69	-0,50	0,60	-0,88
	Membuat kesimpulan	-0,02	0,73	-0,01	0,08	-0,80
Menginterpretasi	Mengidentifikasi kesalahan	0,10	0,70	-0,49	-0,37	-0,88

Untuk lebih jelasnya, dapat dilihat diagram distribusi tingkat kesulitan item menurut aspek dan sub aspek tahap ujicoba berturut-turut adalah menginterpretasi, menganalisis, menginferensi, dan menjelaskan instrumen pada gambar 2. Berdasarkan gambar 2, urutan tingkat kesulitan item masing-masing aspek pada tingkat kesulitan item juga dapat diidentifikasi dari

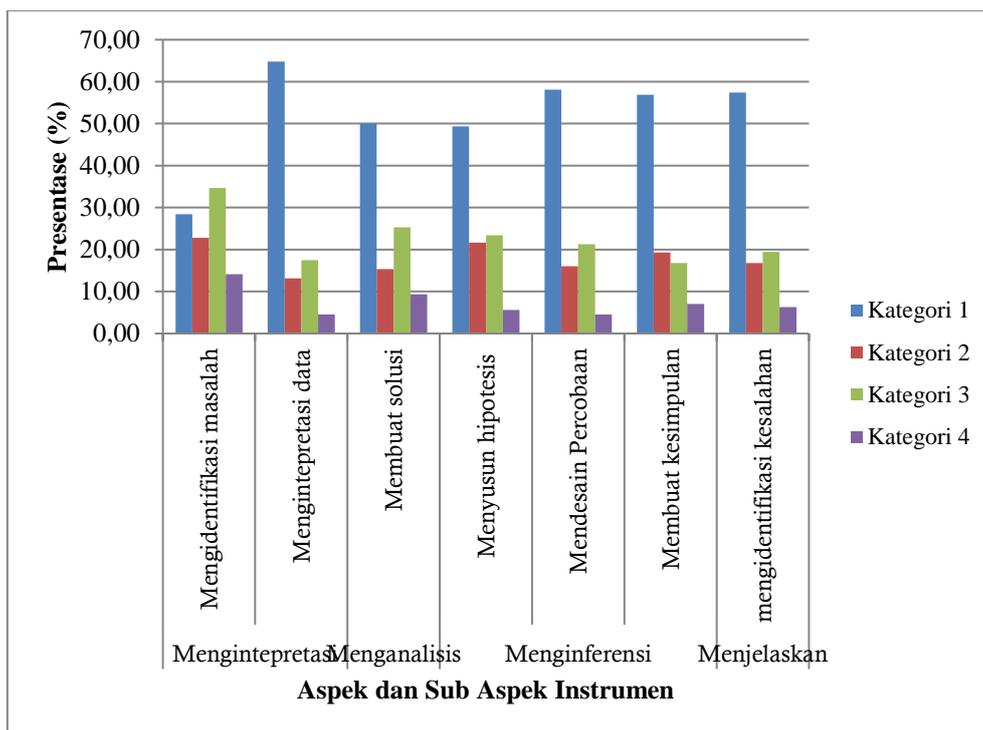
presentase responden menjawab benar untuk masing-masing kategori. Jika presentase responden dapat menjawab kategori 1 paling banyak berarti menggambarkan item tersebut sulit, sebaliknya jika presentase testi paling banyak pada kategori 4 menandakan item tersebut mudah. Presentase responden menjawab benar item pada masing-masing aspek dan sub aspek instrumen untuk keempat kategori dinyatakan dalam tabel 4.



Gambar 2. Tingkat Kesulitan Item Masing-masing Aspek dan Sub Aspek

Tabel 4. Presentase Menjawab Benar Aspek dan Sub Aspek Keterampilan Berpikir Kritis Tahap Ujicoba

Aspek	Sub Aspek	Difficulty	Kesulitan Tahap			
			Kategori 1 (%)	Kategori 2 (%)	Kategori 3 (%)	Kategori 4 (%)
Mengintepretasi	Mengidentifikasi masalah	28,44	22,81	34,64	14,12	28,44
	Mengintepretasi data	64,79	13,17	17,46	4,58	64,79
Menganalisis	Membuat solusi permasalahan	50,00	15,36	25,29	9,35	50,00
Menginferensi	Menyusun hipotesis	49,33	21,66	23,38	5,63	49,33
	Mendesain Percobaan	58,11	16,03	21,28	4,58	58,11
	Membuat kesimpulan	58,87	19,27	16,79	7,06	58,87
Menjelaskan	Mengidentifikasi kesalahan	57,44	16,79	19,47	6,30	57,44



Gambar 3. Presentase Menjawab Benar pada Aspek dan Sub Aspek Masing-masing Kategori Keterampilan Berpikir Kritis pada Tahap Ujicoba

Selanjutnya, berdasarkan tabel 3 tersebut dapat diketahui level keterampilan berpikir kritis peserta didik ditinjau dari aspek dan sub aspek untuk kategori 1, 2, 3, dan 4 yang disajikan pada gambar 3.

Berdasarkan gambar 3 dapat disimpulkan bahwa peserta didik belum menguasai keterampilan berpikir kritis. Hal ini ditunjukkan bahwa sebagian besar peserta didik pada kategori 1 terutama pada sub aspek menganalisis data, sedangkan pada sub aspek mengidentifikasi

masalah, presentase tersebar merata pada tiap kategori yaitu sebesar 28,44% pada kategori 1, 22,81% pada kategori 2, 34,64% pada kategori 3, dan 14,12% pada kategori 4. Urutan sub aspek yang belum dapat dicapai peserta didik selain kedua sub aspek tersebut secara berturut-turut dari presentase kategori 1 yaitu menyusun hipotesis, membuat solusi permasalahan, mengidentifikasi kesalahan, mendesain percobaan, dan membuat kesimpulan.

KESIMPULAN DAN SARAN

Berdasarkan hasil analisis data dan pembahasan dari penelitian yang telah dilakukan maka dapat ditarik simpulan sebagai berikut :

1. Terdapat 50 butir instrumen yang sesuai dengan kriteria model yang berada pada nilai INFITMNSQ mulai dari 0.77 sampai dengan 1.30
2. Tingkat kesukaran rata-rata instrumen yang dikembangkan telah baik dengan

memenuhi kriteria tingkat kesukaran instrumen mulai dari -2 sampai dengan 2 yakni berada pada 0.00 ± 0.24 .

3. Koefisien reliabilitas instrumen yang dikembangkan secara klasik berada pada kategori baik yakni 0.85.
4. Tingkat keterampilan berfikir kritis peserta didik yang paling tinggi berada pada sub aspek menginterpretasi data dan yang

paling rendah berada pada sub aspek mengidentifikasi masalah.

Berdasarkan penelitian yang telah dilaksanakan, beberapa saran untuk penelitian lanjutan yaitu:

1. Instrumen ini tidak digunakan untuk mengukur keterampilan berpikir kritis tetapi untuk memetakan kategori

kemampuan peserta didik saat mengerjakan soal.

2. Reliabilitas data sebaiknya dianalisis dengan memperhatikan SEM (*Standard Error Measurement*) dan ICC (*Item Characteristics Curve*) menggunakan software yang lain.

DAFTAR PUSTAKA

- Adams, R. J. & Khoo, S. T. (1996). *Quest: The interactive test analysis system version 2.1*. Victoria: The Australian Council for Educational Research.
- Dewey, J., Hickman, L. A., & Alexander, T. M. (1998). *The Essential Dewey: Volume 2 Ethics, Logic, Psychology*. Bloomington, IN: Indiana University Press.
- Ennis, R.H. 1996. *Critical Thinking*. USA: Prentice-Hall, Inc.
- Hambleton & Swaminathan. (1985). *Item Response Theory Principles and Applications*. Boston: Kluwer Nijhoff Publishing.
- Johnson, E.B. 2002. *Contextual teaching and learning: What it is and why it's here to stay*. USA: Coewin Press.
- Kementerian Pendidikan dan Kebudayaan. (2013). *Peraturan Kementerian dan Kebudayaan No. 64, Tahun 2013, tentang Standar Isi*.
- Ku, K.Y.L. (2009). Assessing student's critical thinking performance: urging for measurement using multi-response format. *Thinking Skills and Creativity*, 4, 70-76.
- Miri, M., David, B.C., & Uri, Z. (2007). Purposely Teaching for the Promotion of Higher-order Thinking Skills: A Case of Critical Thinking. *Research in Science Education*, 37, 353-369.
- OECD. (2015). *PISA 2015 Scientific Literacy Framework*. Paris: Organization for Economic Cooperation and Development.
- Oriundo, L.L. & Dallo-Antonio, E.M. (1984). *Evaluation Educational Outcomes*. Manila: Rex Printing Compagny, inc.
- Partnership for 21st Century Skills. (2008). 21st Century Skills, Education & Competitiveness. Tersedia di <http://www.p21.org/our-work/p21-framework/> [20-08-2015]
- Wagner, Tony. 2008. *The Global Achievement Gap*. New York: Basic Books.