



PENGEMBANGAN INSTRUMEN TES KEMAMPUAN BERPIKIR TINGKAT TINGGI PADA MATERI HUKUM NEWTON TENTANG GERAK

DEVELOPMENT OF HIGHER THINKING INSTRUMENT OF NEWTON 'S LAWS OF MOTION

Fitri Mar'atus Solekhah*, Nengah Maharta, Wayan Suana
FKIP Universitas Lampung, Jl. Prof. Dr. Soemantri Brojonegoro No.1

*Corresponding author, fitrimaratussolekhah18@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk menghasilkan instrumen tes kemampuan berpikir tingkat materi hukum Newton tentang gerak. Metode yang digunakan ialah metode *research and development*. Model pengembangan mengadopsi dari (Adam dan Wieman, 2011) yang terdiri dari lima tahapan diantaranya menentukan format butir soal, menentukan konstruksi butir soal, menentukan pedoman penilaian, uji ahli, dan revisi butir soal. Pada tahap pengembangan, dilakukan uji ahli kepada tiga dosen yang menguasai materi hukum Newton tentang gerak dan ahli di bidang pengembangan instrument. Subjek uji coba yang digunakan dalam penelitian berjumlah 65 siswa. Data yang telah diperoleh, kemudian dianalisis menggunakan rasch model dengan aplikasi winstep 3.73. Berdasarkan hasil analisis soal, dapat disimpulkan bahwa (1) Instrumen tes kemampuan berpikir tingkat tinggi valid {memenuhi kriteria yang dapat diterima yakni PT-Measure, Outfit Mean Square (MNSQ), dan nilai outfit z- standardized (ZSTD)}, (2) menghasilkan nilai *alpha cronbach* 0,85 yang berarti baik, (3) terdapat 3 soal sulit, 12 soal sedang dan 5 soal mudah, (4) responden memiliki konsistensi dalam menjawab soal yang baik, (5) opsi pengecoh pada semua soal valid, dengan demikian instrumen tes kemampuan berpikir tingkat tinggi materi hukum Newton tentang gerak yang dikembangkan di tingkat sekolah menengah memiliki penguasaan konsep yang baik untuk siswa.

Kata kunci: Kemampuan Berpikir Tingkat Tinggi, Instrumen Tes, Hukum Newton Tentang Gerak, Rasch Model

ABSTRACT

The purpose of this study was to design a test instrument of higher order thinking skills on the Newton's laws of motion. The method used was a research and development. The development model was adapted from Adam and Wieman (2011), consisted of the determination of the instrument format, instrument constructions, assesment guidelines, final test, and revision of the question. At the stage of development, three experts in physics educations assessed the theoretical validity of instrument. After revising the instrument based on the experts suggestion, then it was subjected to 65 students to analyze the instrument's empirical validity. The rasch model using winsteps 3.73 was performed. Based on the results of the analysis, it can be concluded that (1) the test instrument meet the criteria of validity {achieve the accetable ranges of PT-Measure, Outfit Mean Square (MNSQ), and value of z-standardized (ZSTD)}, (2) questions of HOTS developed have excelent reliability with alpha cronbrach's of 0.85, (3) there are 3 difficult questions, 12 medium questions, and 5 easy questions, (4) respondents have good consistency of answer, (5) deception options on all question was valid, thus the test instrument the ability to high level s of thinking of Newton's material laws of motion developed at the secondary school level have a mastery of concepts better to students.

Keywords: The ability to high level of thinking, test instruments, Newton's law of motion, rasch model



1. PENDAHULUAN/ INTRODUCTION

Kemampuan berpikir tingkat tinggi merupakan kemampuan yang penting dimiliki di abad 21 sehingga perlu diajarkan kepada siswa. Hal ini sejalan dengan pendapat Greenhill (2009) yang menyatakan “*Teaching critical thinking in schools is one of the main topics in the discussion regarding so-called 21st Century skills*”. Pada setiap subjek dan pada setiap tingkatan pendidikan, proses pembelajaran perlu mengintegrasikan pembelajaran *content knowledge*, dengan kegiatan-kegiatan yang membentuk kemampuan berpikir tingkat tinggi dan pemecahan masalah (Trilling & Fadel, 2009). Hal ini sesuai dengan versi revisi dari Taksonomi Bloom yang menggunakan istilah *remember, understand, apply, analyze, evaluate* dan *create* (Anderson & Krathwohl, 2001).

Pada dunia kerja abad 21 ini, siswa dituntut untuk memiliki kemampuan berpikir tingkat tinggi agar dapat mengorganisasikan kemampuannya dalam menyelesaikan masalah dengan cara mengevaluasi, menganalisis, dan mencipta atau mengkreasikan suatu gagasan, ide atau suatu konsep yang dapat membangun manusia cerdas dengan intelektual tinggi. Sebab berpikir tingkat tinggi merupakan salah satu tolak ukur tingkat intelektual seseorang. Dimana di kurikulum 2013 ini, pendidikan memiliki peran yang signifikan bahkan menjadi pranata utama dalam menyiapkan sumber daya manusia (Wagiran, 2007).

Kenyataannya, begitu banyak pembelajaran yang terjebak pada kemampuan berpikir tingkat rendah, sedangkan pembelajaran yang lebih mengembangkan keterampilan berpikir tingkat tinggi sangat diperlukan siswa untuk bersaing dan menghadapi tantangan di masa depan. Hal ini sejalan dengan penelitian yang dilakukan Karplus (Mirawati, 2011:4) menunjukkan bahwa masih banyak siswa SMA yang kesulitan dalam menyelesaikan masalah yang memerlukan pemikiran abstrak secara efektif. Selain itu, fakta yang menunjukkan bahwa prestasi fisika yang dimiliki oleh seorang siswa diukur pada aspek *reasoning* Indonesia berada pada ranking 40 dari 42 negara (Micheal & Ina, 2013). Keberhasilan penguasaan suatu konsep akan didapatkan ketika siswa sudah mampu berpikir tingkat tinggi, dimana siswa tidak hanya dapat mengingat dan memahami suatu konsep, namun siswa dapat menganalisis serta mensintesis, mengevaluasi, dan mengkreasikan suatu konsep dengan baik, konsep yang telah dipahami tersebut dapat melekat dalam ingatan siswa dalam waktu yang lama, sehingga penting sekali bagi siswa untuk memiliki kemampuan berpikir tingkat tinggi (Laily, 2013).

Berdasarkan pernyataan tersebut, nyatanya di salah satu SMA yang ada di Bandar Lampung belum ada instrumen khusus yang digunakan untuk mengukur kemampuan berpikir tingkat tinggi siswa. Hal ini dapat diketahui dari data analisis lapangan terkait jenis instrumen yang digunakan oleh guru fisika di sekolah. Instrumen yang sering digunakan guru dalam pembelajaran di sekolah adalah *traditional assessment* (tes tulis) dalam bentuk pilihan ganda (*multiple choice*). Instrumen evaluasi yang mengukur kemampuan berpikir tingkat tinggi dapat menggunakan berbagai tipe penilaian seperti *modified multiple choice*, konstruksi jawaban singkat, dan konstruksi jawaban panjang seperti yang telah dilakukan oleh Ramirez dan Ganaden (2008). Salah satu alternatif *Modified multiple choice* yang dapat digunakan untuk mengukur kemampuan berpikir tingkat tinggi adalah bentuk *two-tier multiple choice question* (pilihan ganda bertingkat). Bentuk soal *two-tier multiple choice question* dikembangkan oleh Treagust (2006). Bentuk soal *two-tier multiple choice question* dapat digunakan untuk membantu menguji pemahaman siswa serta membantu mengidentifikasi konsep yang mungkin dimiliki oleh siswa. Cullinane (2011) mengemukakan penyertaan alasan pada tingkatan kedua dari bentuk soal *two-tier multiple choice question* dapat digunakan untuk meningkatkan keterampilan berpikir tingkat tinggi dan melihat kemampuan siswa dalam memberi alasan. Penyertaan alasan pada tingkatan kedua soal ini dapat digunakan untuk mengurangi terjadinya untung-untungan yang sering menjadi kelemahan dari bentuk soal pilihan ganda biasa.

Penilaian dalam pendidikan menggunakan dua macam teori pengukuran, yakni: teori pengukuran klasik dan teori pengukuran modern. Menurut Sumintono & Widhiarso (2014) keunggulan pemodelan Rasch dibanding metode lainnya, khususnya teori tes klasik, yaitu kemampuan melakukan prediksi terhadap data hilang (*missing data*), berdasarkan pola respon individu. Keunggulan ini menjadikan hasil analisis statistik model Rasch lebih akurat dalam penelitian yang dilakukan, dan yang lebih penting lagi, pemodelan Rasch mampu menghasilkan nilai pengukuran eror standar untuk instrumen yang digunakan dapat meningkatkan ketepatan perhitungan. Kalibrasi



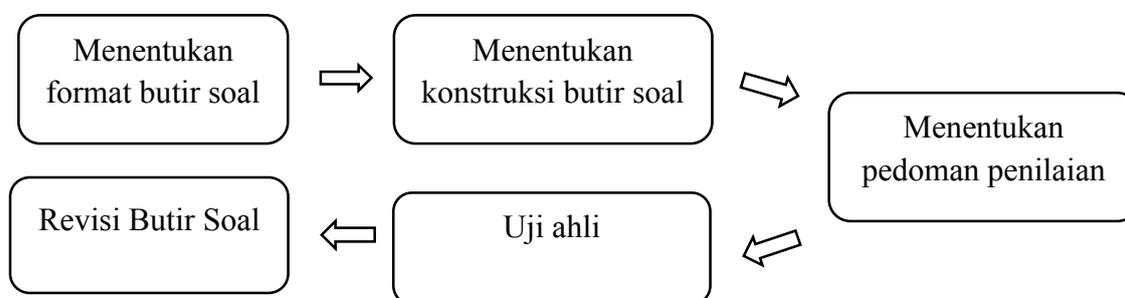
dilakukan dalam pemodelan Rasch secara sekaligus dalam tiga hal, yaitu skala pengukuran, responden (*person*), dan butir soal.

Berdasarkan uraian tersebut, untuk mengukur kemampuan berpikir tingkat tinggi materi hukum Newton tentang gerak digunakan tes pilihan ganda beralasan atau *two-tier multiple choicequestion*. Untuk itu perlu disusun instrumen penilaian kemampuan berpikir tingkat tinggi Fisika yang terdiri atas tes dan pedoman penilaian. Dengan demikian, diperlukan adanya analisis model rasch pada instrumen tes kemampuan berpikir tingkat tinggi materi hukum newton tentang gerak.

2. METODE PENELITIAN/ RESEARCH METHODE

Penelitian ini menggunakan metode penelitian dan pengembangan (*Research and Development*) menurut (Adams & Wieman, 2011). Penelitian ini dilakukan di salah satu SMA yang ada di Bandar Lampung dengan sampel sebanyak 65 siswa. Pengumpulan data yangdigunakan pada penelitian ini menggunakan teknik skala untuk siswa.Teknik skala merupakan pengumpulan data yang dilakukandengan cara memberi seperangkat pertanyaan tertulis kepada responden untukdijawab (Sugiyono, 2012).Selain teknik skala, teknik wawancara juga dilakukan dalam penelitian ini untuk mengumpulkan data-data mengenai pembelajaran pada materi hukum Newton tentang gerak, serta instrumen tes yang digunakan selama proses pembelajaran, teknik wawancara dilakukan kepada salah satu guru fisika kelas XI SMA Negeri 05 Bandar Lampung.

Penelitian ini dilakukan di salah satu SMA yang ada di Bandar Lampung dengan 32 orang siswa kelas XI IPA. Selanjutnya, dilakukan uji ahli kepada tiga dosen, uji keterbacaan kepada empat siswa yang dipilih secara random dan uji terbatas kepada 65 siswa. Adapun dalam pelaksanaannya, penskoran yang digunakan adalah *Graded Response Model* (GRM) karena data yangdihasilkan berupa data politomus. Analisis data ini menggunakan aplikasi winsteps 3.73 yang memenuhi nilai outfit mean square, nilai outfit Z-standard (ZSTD) dan nilai Pt Mean Core. Pendekatan model Rasch ini, selain memperhatikan butir soal juga memerhatikan aspek responden. Secara lebih rinci alur pengembangan instrumen tes kemampuan berpikir tingkat tinggi dapat dilihat pada Gambar1.



Gambar 1. Alur Pengembangan Instrumen Tes Kemampuan Berpikir Tingkat Tinggi

3. HASIL DAN PEMBAHASAN/ RESULT AND DISCUSSION

3.1 Hasil Penelitian

3.1.1 Hasil Uji Ahli

Pengembangan soal berpikir tingkat tinggi materi hukum Newton tentang gerak ini sebanyak 20 soal yakni soal pilihan ganda beralasan dengan 8 soal menganalisis (C4), 6 soalmengevaluasi (C5), dan 6 soal mencipta (C6). Soal berpikir tingkat tinggi yang telah dikembangkan kemudian diujikan kepada tiga dosen FKIP Pendidikan Fisika yang menguasai materi hukum Newton tentang gerak dan ahli dibidang pengembangan instrumen. Hal ini sejalan dengan pendapat Gelerstein *et al*, (2016) yang menyatakan bahwa uji ahli dilakukan untuk memeriksa kesesuaian hubungan antara pertanyaan-pertanyaan dengan tujuan tes.

Nilai validasi yang diperoleh sebesar 81,24% dengan kriteria valid, instrumen tes dapat digunakan dengan revisi kecil. Berdasarkan rata-rata hasil dari tiga dosen uji ahli diperoleh bahwa



validitas instrumen dari aspek materi sebesar 83,33% yang dikategorikan sesuai dengan kriteria validitas instrumen maka instrumen tes tersebut dinyatakan valid (boleh digunakan dengan revisi kecil), aspek konstruksi sebesar 80,56% yang dikategorikan sesuai dengan kriteria validitas instrumen maka instrumen tes tersebut dinyatakan valid (boleh digunakan dengan revisi kecil), dan aspek bahasa sebesar 79,83% yang dikategorikan sesuai dengan kriteria validitas instrumen maka instrumen tes tersebut dinyatakan sangat valid (boleh digunakan dengan revisi kecil). Hasil validasi ahli secara keseluruhan dapat dilihat pada Tabel 1.

Tabel 1. Hasil Validasi Ahli Secara Keseluruhan

Ahli	Nilai		
	Materi	Konstruk	Bahasa
Ahli 1	81	81	81
Ahli 2	82	82	82
Ahli 3	87	80,67	76,5
Nilai Rata-rata	83,33	80,23	79,83

Berdasarkan penilaian validasi instrumen secara menyeluruh dari tiga dosen ahli, maka instrumen tes berpikir tingkat tinggi pada materi hukum Newton tentang gerak yang dikembangkan dinyatakan layak digunakan untuk mengukur kemampuan berpikir tingkat tinggi siswa dengan revisi kecil.

3.1.2 Hasil Uji Keterbacaan

Instrumen tes yang telah divalidasi oleh tiga dosen ahli dan dinyatakan valid serta layak digunakan, kemudian diuji keterbacaannya oleh empat orang siswa yang dipilih dengan cara random untuk melihat apakah siswa telah paham dengan maksud pertanyaan yang terdapat pada butir-butir soal yang diujikan. Empat orang siswa tersebut ialah dua siswa dari kelas XI IPA 1 dan dua siswa dari kelas XI IPA 3. Instrumen uji coba keterbacaan ini menggunakan teknik *scaling* yang terdiri dari beberapa pertanyaan dengan respon jawaban siswa “Ya” jika jawaban dari pertanyaan sesuai dengan karakteristik instrumen tes dan “Tidak” jika jawaban dari pertanyaan tidak sesuai dengan karakteristik instrumen tes. Pertanyaan yang terdapat pada angket yaitu mengenai kejelasan bahasa, huruf, angka, gambar, dan grafik yang terdapat dalam soal.

Hasil dari uji keterbacaan oleh empat orang siswa menunjukkan bahwa instrumen tes kemampuan berpikir tingkat tinggi materi hukum Newton tentang gerak yang dikembangkan telah jelas. Artinya, bahasa yang digunakan pada butir-butir soal yang dikembangkan dapat dipahami dan siswa tidak mengalami penafsiran ganda. Oleh karena itu, instrumen tes kemampuan berpikir tinggi yang dikembangkan dinyatakan layak dan dapat digunakan dalam uji coba kelas terbatas.

3.1.3 Hasil Uji Coba Terbatas

Uji coba terbatas ini dilakukan di salah satu SMA yang ada di Bandar Lampung dengan jumlah 65 siswa. Instrumen tes yang diujikan berupa soal pilihan ganda beralasan sebanyak 20 butir soal yang terdiri dari 8 soal menganalisis (C4), 6 soal mengevaluasi (C5), dan 6 soal mencipta (C6). Soal-soal tersebut, sebelum diujikan kepada siswa diuji cobakan terlebih dahulu dari segi bahasa, konstruksi, dan materi oleh tiga orang dosen ahli dan uji keterbacaan oleh empat orang siswa SMA yang telah mempelajari materi hukum Newton tentang gerak. Selanjutnya, butir soal di revisi dan dinyatakan valid oleh tiga orang dosen ahli dan layak digunakan, maka instrumen tes diujicobakan kepada 65 responden. Hasil dari instrumen tes kemampuan berpikir tingkat tinggi tersebut yang dikerjakan oleh 65 siswa diuji ke validan dan reliabilitas menggunakan model *rasch* dengan aplikasi *winsteps 3.73* dengan cara data yang telah diperoleh, ditabulasi dalam *software Ms. Excel* untuk kemudian dikonversikan dan dianalisis dengan bantuan *software winsteps 3.73* dalam sistem operasi *windows 7*.

Pada analisis tahap pertama, peneliti mengidentifikasi butir soal atau *item* yang berjumlah 20 butir soal. Tingkat kesesuaian butir soal (*item fit*) digunakan untuk menjelaskan apakah butir soal berfungsi normal melakukan pengukuran atau tidak. Apabila didapati suatu soal yang tidak fit, berarti



terdapat indikasi bahwa terjadi miskonsepsi pada siswa terhadap butir soal tersebut. Menurut Boone, dkk. (2014), parameter yang digunakan untuk mengetahui kesesuaian responden antara lain: Pertama, nilai *outfit mean square* yang diterima $0,5 < \text{MNSQ} < 1,5$. Kedua, nilai *outfit Z-standard* yang diterima: $-2,0 < \text{ZSTD} < +2,0$. Ketiga, nilai (Pt Mean Corr) diterima: $0,4 < \text{Pt Measure Corr} < 0,85$. Berdasarkan ketiga kriteria tersebut, semua butir soal telah memenuhi kriteria sehingga dapat disimpulkan bahwa semua soal valid dan tidak perlu diubah atau diganti. Hasil analisis akhir dari pengembangan instrumen berpikir tingkat tinggi materi hukum Newton tentang gerak yang berisi 20 butir soal dengan jumlah responden 65 responden tersaji dalam Tabel 2.

Tabel. 2 Rangkuman Hasil Analisis Akhir

	Output	Hasil
Butir Soal	Reliabilitas butir soal	0,86
	Indeks Separasi	2,49
	Pemisahan Strata	3,65
	Nilai Logit Tertinggi	0,46logit (S15)
	Nilai Logit terendah	-1,83logit (S1)
Responden	Reliabilitas responden	0,83
	Indeks Separasi	2,18
	Pemisahan Strata	3,24
	Nilai Logit Tertinggi	2,76logit (35L)
	Nilai Logit Terendah	-2,26 logit (39P)
Instrumen	<i>Alpha Cronbrach</i>	0,85

Hasil analisis data diperoleh bahwa 20 soal dinyatakan valid dengan nilai sebesar 0,86 menunjukkan bahwa kualitas butir soal dalam instrumen ini tergolong tinggi. Sehingga, keduapuluh butir soal yang teridentifikasi tersebut, memiliki ketepatan butir-butir soal yang berkualitas. Selanjutnya, nilai reliabilitas responden sebesar 0,83 menunjukkan bahwa konsistensi jawaban dari responden tergolong tinggi. Artinya, responden menjawab keseluruhan butir-butir soal dengan sungguh-sungguh dan tidak asal-asalan. Nilai *alpha cronbach* yang diperoleh dari hasil analisis instrumen adalah sebesar 0,85. Artinya, reliabilitas yang diperoleh alat ukur ini tergolong bagus. Sehingga instrumen tes yang digunakan sudah dapat mengukur apa yang hendak diukur.

Berdasarkan tabel 2, diketahui bahwa butir soal yang paling susah dikerjakan oleh responden adalah butir soal bernomor 15(S15) dengan nilai 0,46 logit yang merupakan nilai tertinggi di antara butir-butir soal yang lain. Sementara itu, aitem yang paling mudah dikerjakan oleh responden adalah butir soal bernomor satu (S1) dengan nilai -1,83 logit yang merupakan nilai paling rendah. Dengan kata lain, butir soal bernomor 15 (S15) memiliki tingkat kesulitan tertinggi (0,46 logit) dan butir soal bernomor 1(S1) memiliki tingkat kesulitan terendah (-1,83 logit).

Responden yang mempunyai kemampuan berpikir tingkat tinggi adalah responden nomor 35 dengan jenis kelamin laki-laki dengan nilai 2,76 logit yang merupakan nilai logit tertinggi. Sebaliknya, responden yang mempunyai kemampuan berpikir tingkat rendah adalah responden nomor 39 dengan jenis kelamin perempuan, ditunjukkan dari nilai -2,26 logit yang merupakan nilai logit terendah. Pengelompokkan responden dan butir soal dapat diketahui dari indeks separasi. Indeks separasi responden yang diperoleh adalah 2,18 dan indeks separasi aitemnya adalah 2,49. Untuk melihat pengelompokkan secara lebih detail dapat digunakan persamaan pemisahan strata $H = \{(4 \times \text{indeks separasi}) + 1\}/3$. Hasilnya, untuk responden $H = 3,24$ (dibulatkan jadi 3). Sementara itu, untuk butir soal $H = 3,65$ (dibulatkan jadi 4). Nilai indeks separasi dan pemisahan strata baik pada butir soal dan responden tergolong besar. Hal ini menunjukkan bahwa kualitas soal dalam keseluruhan responden dan butir soal makin bagus karena mampu mengidentifikasi kelompok responden dan kelompok aitem secara heterogen.

Setelah dilakukan analisis menggunakan model *Rasch*, diperoleh 20 butir soal yang memiliki indeks ketepatan model sesuai kriteria. Selanjutnya dilakukan analisis *rating scale* yang bertujuan untuk memverifikasi apakah peringkat pilihan yang digunakan membingungkan responden atau tidak. Hasil analisisnya tersaji dalam Tabel 3.



Tabel 3. *Rating Scale*

OBSVD AVRGE	ANDRICH TRESHOLD	Score
-.47	NONE	0
.45	-1.09	1
1.47	1.09	2

Pada tabel 3 menunjukkan bahwa opsi yang diberikan sudah valid bagi responden karena nilai *logit* pada kolom OBSVD AVRGE menunjukkan bahwa nilai *logit* yang meningkat dari rendah sampai tinggi dan nilai *Andrich Threshold* bergerak dari NONE kemudian negatif dan terus mengarah ke positif secara berurutan. Secara rinci instrumen tes mengenai analisis validitas, reliabilitas, dan *rating score* soal dapat dilihat pada lampiran.

Setelah dilakukan analisis menggunakan model *Rasch*, diperoleh 20 aitem yang memiliki indeks ketepatan model sesuai kriteria. Untuk mengetahui tingkat kesulitan butir soal (*item measure*) dilihat dari nilai *logit* tiap butir soal yang dapat dilihat pada kolom *measure*. Nilai *logit* yang tinggi menunjukkan tingkat kesulitan soal yang paling tinggi. Suminto dan Widhiarso (2015) memberikan panduan dalam menilai butir soal tersebut menjadi empat kategori yang terlihat dalam Tabel 4.

Tabel 4. Hasil Tingkatan Butir Soal Kemampuan Berpikir Tingkat Tinggi Siswa

Kategori	Nilai Measure	No Soal	Jumlah Soal
Sulit	0 sd 1	15,20,dan 19	3
Sedang	-1 sd 0	3,4,5,9,10,11,12,13,14,16,17,dan18	12
Mudah	< -1	1,2,7,6,dan 8	5
	Total	20	20

Ada beberapa hal yang perlu diperhatikan dari output analisis *Rasch* dengan *wisntep 3.73* ini. Nilai *logit (measure)* yang tinggi menunjukkan bahwa butir soal tersebut memiliki tingkat kesulitan yang tinggi. Hal ini berkorelasi dengan total *score*, dimana jumlah penjawab benar yang sedikit dalam total *score* berkorelasi dengan nilai *measure* yang semakin tinggi.

3.1.4 Hasil Kemampuan Berpikir Tingkat Tinggi Siswa

Instrumen tes kemampuan berpikir tingkat tinggi diujicobakan sebanyak 20 butir soal kepada 65 responden. Skor-skor yang diperoleh dari hasil uji coba kemudian diolah menjadi data. Data ditabulasi dalam *software ms. Exel* untuk kemudian dikonversikan dan dianalisis dengan bantuan *software winstep 3.73* dalam sistem operasi *windows 7*.

Kemampuan berpikir tingkat tinggi siswa dari hasil uji dapat dilihat di *person statistics: measure order* yang terdapat di lampiran. Semakin tinggi nilai *measure* maka kemampuan berpikir siswa semakin tinggi. Selain itu, kemampuan berpikir tingkat tinggi siswa juga dapat dilihat pada *person map item*. Pada *person map item* bagian kanan menunjukkan tingkat kesulitan soal sedangkan bagian kiri menunjukkan kemampuan berpikir tingkat tinggi. Urutan kemampuan berpikir tingkat tinggi dan jenis soal yang dianggap sulit oleh siswa diurutkan dari atas ke bawah.

3.2 PEMBAHASAN

Indikator-indikator yang digunakan untuk mengembangkan instrumen tes kemampuan berpikir tingkat tinggi materi hukum Newton tentang gerak ialah stimulus yang berupa wacana, grafik, ataupun gambar. Indikator-Indikator pada instrumen tes kemampuan berpikir tingkat tinggi ini menggunakan Kata Kerja Operasional (KKO) Anderson dan Krathworl yang disesuaikan dengan level kognitif yang diukur, sehingga pada penelitian ini menggunakan indikator soal instrumen tes kemampuan berpikir tingkat tinggi yang sesuai dengan taksonomi Bloom revisi yang mengukur kemampuan berpikir tingkat tinggi siswa, yaitu kemampuan menganalisis (C4), mengevaluasi (C5), dan mencipta (C6).

Tahap yang dilakukan dalam pengembangan indikator-indikator instrument tes berpikir tingkat tinggi materi hukum Newtn tentang gerak berdasarkan KD 3.4, dan 4.4 yang mengacu pada kurikulum 2013 revisi yaitu menentukan materi apa yang akan diukur dan level kognitif yang dapat mengukur materi tersebut, dan menentukan stimulus yang tepat digunakan untuk mengembangkan KD tersebut.



Instrumen yang dikembangkan memiliki karakteristik instrumen kemampuan berpikir tingkat tinggi, dimana instrumen tes tersebut menggunakan informasi untuk menyelesaikan masalah, menelaah informasi secara kritis, mentransfer satu konsep ke konsep lain, mencari kaitan dari informasi yang berbeda. Instrumen tes berupa soal tersebut mengukur kemampuan berpikir tingkat tinggi dengan cara menampilkan permasalahan yang kontekstual atau realistis dalam kehidupan sehari-hari, namun masih tetap mengaitkan dengan konsep yang dipelajari.

Spesifikasi instrumen tes berpikir tingkat tinggi materi hukum Newton tentang gerak fisika SMA yang dikembangkan yaitu:

- a. Instrumen yang dikembangkan mengacu pada indikator kemampuan berpikir tingkat tinggi berpikir tingkat tinggi menurut Anderson dan Krathworl dengan menggunakan Kata Kerja Operasional (KKO) yang mengacu pada taksonomi bloom yang telah direvisi.
- b. Tujuan dari pengembangan instrumen tes berpikir tingkat tinggi adalah menghasilkan instrumen tes berpikir tinggi khususnya pada materi hukum Newton tentang gerak yang valid dan reliabel sehingga dapat digunakan oleh guru untuk mengukur kemampuan berpikir tingkat tinggi siswa
- c. Soal yang dikembangkan mematuhi tingkat kognitif dari kemampuan menganalisis (C4), mengevaluasi (C5), dan mencipta (C6)
- d. Kata kerja operasional dalam indikator yang digunakan untuk mengembangkan soal HOTS hukum Newton tentang gerak ini antara lain untuk C4 (menganalisis, mendeteksi, dan menyimpulkan, C5 (memprediksi, mengkritisi) dan C6 (memberikan solusi, rancangan, serta mengusulkan ide).
- e. Jenis instrumen tes yang dikembangkan adalah tes pilihan ganda beralasan dengan lima pilihan jawaban.
- f. Siswa memilih jawaban yang dianggapnya benar lalu memberikan alasan atas jawaban tersebut pada kolom yang tersedia
- g. Waktu pengerjaan 20 butir soal instrumen tes berpikir tingkat tinggi 100 menit
- h. Soal dikerjakan sendiri dan tutup buku.

Instrumen tes kemampuan berpikir tingkat tinggi yang dikembangkan telah memenuhi standar untuk penilaian, karena instrumen tersebut memiliki reliabilitas, validitas, serta *rating scale* yang baik. Dimana soal-soal kemampuan berpikir tingkat tinggi yang dikembangkan tersebut merupakan soal yang kontekstual dan menggambarkan kejadian kejadian yang nyata (realistis) namun tetap mengaitkan pada konsep pembelajaran yang telah dipelajari.

Instrumen berpikir tingkat tinggi yang telah dikembangkan dan telah tervalidasi oleh tiga dosen ahli kemudian diujikan kepada 65 orang siswa dengan rincian 32 siswa kelas IPA 1 dan 33 siswa kelas IPA 2 untuk dianalisis validitas dan reliabilitasnya. Setelah dilakukan analisis terhadap instrumen secara keseluruhan yang telah diujikan dan direvisi, maka akan diperoleh hasil validitas dan reliabilitas instrumen tes berpikir tingkat tinggi.

Proses pembuatan instrumen tes kemampuan berpikir tingkat tinggi melalui beberapa tahapan penelitian dan pengembangan. Tahapan yang digunakan dalam mengembangkan produk ini mengacu pada prosedur penelitian dan pengembangan instrumen tes menurut (Adams & Wieman, 2011) yang terdiri dari lima tahap yaitu menentukan format butir soal, menentukan konstruksi butir soal, menentukan pedoman penilaian, uji ahli kepada tiga dosen yang ahli dalam bidangnya, dan revisi butir soal, sebagaimana yang telah dipaparkan pada bagian hasil pengembangan produk dalam bab III.

Instrumen tes keterampilan berpikir tingkat tinggi yang dikembangkan mengacu pada Anderson dan Krathworl (2001) yang disesuaikan dengan dimensi kognitif dan dimensi pengetahuan Taksonomi Bloom revisi. Indikator yang digunakan pada dimensi kognitif untuk mengukur kemampuan berpikir tingkat tinggi siswadiantaranya menganalisis (C4), mengevaluasi (C5), dan mencipta (C6). Kata Kerja Operasional yang digunakan untuk masing-masing indikator dimensi kognitif antara lain C4 (menganalisis, mendeteksi dan menyimpulkan), C5 (memprediksi dan mengkritisi), dan C6 (mengusulkan ide dan memberikan rancangan). Selain itu, indikator yang digunakan pada dimensi pengetahuan untuk mengukur kemampuan berpikir tingkat tinggi siswadiantaranya pengetahuan faktual (PF), pengetahuan konseptual (PK), pengetahuan Metakognitif, dan pengetahuan prosedural (PP). Bentuk instrumen penilaian kemampuan berpikir tingkat tinggi yang dikembangkan adalah tes pilihan ganda beralasan dengan lima pilihan jawaban. Butir-butir soal yang telah dikembangkan pada penelitian ini sebanyak 20 soal pilihan ganda beralasan pada materi hukum Newton tentang gerak yang terdiri dari 8 soal menganalisis (C4), 6 soal mengevaluasi (C5), dan 6 soal mencipta (C6).



Instrumen tes yang dikembangkan, sebelum diujicobakan kepada 65 orang siswa diujicobakan kevalidannya terlebih dahulu kepada ahli dari segi aspek konten atau materi, konstruk, dan bahasa agar data yang diperoleh nantinya dapat dipertanggungjawabkan. Hasil validasi ahli dari instrumen tes yang dikembangkan dari aspek materi sebesar 83,33%, aspek konstruksi atau konten sebesar 80,56%, dan aspek materi sebesar 79,83% serta secara keseluruhan diperoleh nilai hasil validasi ahli dari ketiga aspek yaitu sebesar 81,24% yang dikategorikan pada kriteria hasil berpikir tinggi termasuk kedalam instrumen tes yang valid dan layak dengan revisi kecil.

Revisi yang dilakukan yaitu revisi kecil karena sesuai dengan hasil uji validasi, instrumen tersebut dinyatakan valid dan boleh digunakan namun dengan melakukan revisi kecil. Revisi tersebut diantaranya, yaitu:

a. Revisi produk dari segi aspek bahasa

Setelah melakukan validasi terhadap instrumen tes berpikir tingkat tinggi yang dikembangkan, instrumen dinyatakan valid namun harus direvisi sesuai dengan saran validator. Adapun saran dari validator yaitu: perbaiki penulisan yang harus disesuaikan dengan EYD, perbaiki tata bahasa dan susunan kalimat agar lebih mudah dipahami.

b. Revisi produk dari segi aspek konstruksi

Kalimat pada soal no 14 yang salah persepsi sehingga perlu diperbaiki. Pilihan jawaban soal no 3 yang kurang sesuai sehingga harus diperbaiki, serta gambar no 1 dan 15 belum begitu jelas sehingga harus diperbaiki.

c. Revisi produk dari segi aspek konten/materi

Soal nomor 10 diganti karena belum sesuai dengan kondisi di lingkungan sekitar. Selain itu, kisi-kisi instrumen terdapat beberapa catatan kecil seperti kalimat-kalimat yang terdapat pada indikator harus diperhatikan untuk menambah kejelasan soal.

Revisi validasi secara umum adalah pada tata tulis dalam butir soal, kesesuaian indikator, ranah kognitif soal, ketepatan gambar dan istilah. Hasil validasi secara keseluruhan dari aspek bahasa, konstruksi, dan materi diperoleh sebesar 81,46% sehingga instrumen tes berpikir tingkat tinggi dinyatakan valid dan dapat digunakan uji coba terbatas setelah dilakukan revisi.

Instrumen tes yang tervalidasi oleh tiga dosen ahli tadi selanjutnya diuji keterbacaannya oleh empat orang siswa SMA yang diambil secara random dengan tujuan ketika uji coba terbatas, siswa tidak mengalami penafsiran ganda dari pertanyaan dalam tiap-tiap butir soal yang berkaitan dengan penggunaan bahasa dalam soal yang mungkin sulit dipahami, kejelasan huruf, angka, gambar, dan grafik yang disajikan dalam soal. Setelah uji keterbacaan, kemudian dilakukan uji coba terbatas kepada 65 siswa kelas XI IPA yakni kelas XI IPA 1 sebanyak 32 siswa dan kelas XI IPA3 sebanyak 33 siswa di SMA N 05 Bandar Lampung. Uji coba terbatas ini dilakukan pada tanggal 05 Mei 2018 dan tanggal

07 Mei 2018.

Analisis data dari hasil uji coba menggunakan *winsteps* 3.73 diperoleh 20 soal dinyatakan valid karena telah memenuhi ketiga kriteria dalam pemodelan *rasch model* yaitu:

Pertama, nilai *outfit mean square* yang diterima $0,5 < MNSQ < 1,5$. Kedua, nilai *outfit Z-standard* yang diterima: $-2,0 < ZSTD < + 2,0$. Ketiga, nilai *Pt Mean Corr* diterima: $0,4 < Pt Measure Corr < 0,85$.

Selain itu, soal juga dinyatakan reliabel dengan nilai *alpha cronbach's* sebesar 0,89 yang termasuk kedalam kategori instrumen dengan reliabilitas bagus sekali. Berdasarkan hasil uji coba instrumen tes, sebanyak 20 soal yang dikembangkan memiliki tingkat kesulitan yang berbeda-beda diantaranya 2 soal sangat sulit, 9 soal sulit, 6 soal mudah dan 3 soal sangat mudah. Hasil penelitian di SMA N 05 Bandar Lampung yang terdiri dari 65 responden memiliki nilai indeks separasi dan pemisahan strata baik pada butir soal dan responden tergolong cukup besar. Hal ini menunjukkan bahwa skala yang dihasilkan memiliki kualitas yang baik, karena mampu mengidentifikasi kelompok responden secara heterogen dan aitem dengan cukup teliti. Kemampuan berpikir tingkat tinggi siswa yang dapat dilihat pada lampiran (*person statistics: measure order* dan *peta wright*).

Soal pilihan ganda beralasandapat dijadikan alternatif untuk mengukur kemampuan berpikir tingkat tinggi siswa pada materi hukum Newton tentang gerak. Hal ini sesuai dengan hasil penelitian Shidiq, dkk. (2014), yang menyatakan bahwa instrumen tes pilih ganda beralasan dapat digunakan untuk mengukur kemampuan berpikir tingkat tinggi siswa. Selain itu, menurut Bond dan Fox (2007)



menyampaikan bahwa penggunaan model Rasch dalam validasi instrumen akan menghasilkan informasi yang lebih holistik tentang instrumen dan lebih memenuhi definisi pengukuran. Oleh karenanya, penelitian ini pun menggunakan pemodelan Rasch dalam analisis datanya. Untuk memperjelas hasil analisis model Rasch dalam penelitian ini, dilakukan pula pengujian reliabilitasnya menggunakan pendekatan teori tes klasik. Hal ini dilakukan untuk mengetahui hasil analisis model Rasch terhadap instrument tes kemampuan berpikir tingkat tinggi siswa SMA.

4. KESIMPULAN DAN SARAN/ CONCLUSION

4.1 Kesimpulan

Berdasarkan data hasil penelitian dan pembahasan dapat disimpulkan bahwa (1) indikator yang digunakan untuk mengembangkan instrumen tes kemampuan berpikir tingkat tinggi materi hukum Newton tentang gerak menggunakan dimensi kognitif dan dimensi pengetahuan Taksonomi Bloom yang telah direvisi, (2) instrumen tes berpikir tingkat tinggi materi hukum Newton tentang gerak yang dikembangkan telah memenuhi standar kelayakan instrumen yaitu valid dan reliabel. Hasil analisis data uji coba menggunakan model *rasch model* dengan aplikasi *winsteps 3,73* telah memenuhi kriteria kesesuaian butir soal bahwa semua butir soal valid, (3) instrumen tes berpikir tingkat tinggi yang dikembangkan memiliki reliabilitas tinggi dengan nilai *alpha cronbach* sebesar 0,85 yang berarti bagus sekali, (4) responden memiliki konsistensi dalam menjawab soal yang baik, (5) opsi pengecoh pada semua soal valid. Dengan demikian, instrumen tes berpikir tingkat tinggi pada materi hukum Newton tentang gerak dapat digunakan sebagai alternatif untuk mengukur kemampuan berpikir tingkat tinggi siswa.

4.2 Saran

Saran dari penelitian pengembangan ini sebagai berikut:

1. Bagi guru, diharapkan guru dapat mengembangkan instrumen tes berpikir tingkat tinggi berdasarkan indikator berpikir tingkat tinggi pada setiap materi fisika. Dengan demikian, guru dapat melatih kemampuan berpikir tingkat tinggi siswa.
2. Bagi sekolah, hasil penelitian diharapkan dapat menjadi salah satu dasar untuk mengetahui tingkat kemampuan berpikir tingkat tinggi siswa, sehingga pihak sekolah dapat memfasilitasi guru untuk mengembangkan instrumen tes berpikir tingkat tinggi.
3. Bagi peneliti, diharapkan peneliti berikutnya dapat melakukan penelitian mengenai pengembangan instrumen kemampuan tes berpikir tinggi dengan indikator dimensi pengetahuan yang lebih lengkap.

5. DAFTAR PUSTAKA/ REFERENCES

1. Adams, W. K. & Wieman, C. E. 2011. Development and validation of instruments to measure learning of expert like thinking. *International Journal of Science Education*.
2. Anderson, L.W., dan Krathwohl, D.R. 2001. A Taxonomy for Learning, Teaching, and Assesing: A Revision of Bloom's Taxonom y of Educatioanl Objectives. New York: Addison Wesley Longman, In.
3. Bernie Trilling & Charles Fadel. 2009. 21st Century Skills Learning for Life in Our Times. San Fransisco: Wiley.
4. Bond, T.G., & Fox, C. 2007. *Applying the rasch model. Fundamental measurement in the human sciences*. New Jersey : Lawrence Erlbaum.
5. Boone, W.J., Staver, J.R., dan Yale, M.S. 2014. *Rasch Analysis in The Human Sciences*. Dordrecht: Springer.
6. Centeno, E. G. & Sompong, N. 2012. Technology and Learning Styles in the GE Classroom: Towards Developing Blended Learning Systems for the 21 th Century Learner. *British Journal of Arts and Social Sciences*.



7. Cullinane, Alison dan Maeve Liston. 2011. *Two-tier Multiple Choice Question: An Alternative Method of Formative Assessment for First Year Undergraduate Biology Students*. Limerick: National Center for Excellence In Mathematics and Education Science Teaching and Learning (NCE-MSTL).
8. Gelerstein, D., Rio, R. D., Nussbaum, M., Chiuminatto, P., & Lopez, X. 2016. Designing and implementing a test for measuring critical thinking in primary school. *Science Direct*, Thinking Skill and Creative. Available <http://www.elsevier.com/locate/tsc>.
9. Greenhill. 2009. *The Physics of Everyday Phenomena: A Conceptual Introduction to Physics, Sixth Edition*. New York: Mc Graw-Hill Companies.
10. Ima, Michael. 2013. *Michael Allen's Guide to E-learning*. Canada.
11. Laily, Nur Rochmah. 2013. Analisis Soal Tipe *Higher Order Thinking Skill* (HOTS). *Jurnal unswagati*. Tersedia di jurnal.unswagati.ac.id/index.php/Euclid/article/download/323/203 (online). Diakses pada 20 April 2017. 9 (1).
12. Mirawati, Riani. 2011. Penerapan Model Pembelajaran STAD untuk Meningkatkan Kemampuan penalaran Ilmiah dan Pemahaman Konsep Siswa SMA pada Konsep Pencernaan. Skripsi Sarjana UPI. Bandung: Tidak diterbitkan.
13. Misbah, I.H., & Sumintono, B. 2014. Pengembangan dan Validasi Instrument "Persepsi Siswa Terhadap Karakter Moral Guru" di Indonesia Dengan Model Rasch. Dipresentasikan dalam Seminar Nasional "Pengembangan Instrumen Penilaian Karakter yang Valid" di Fakultas Psikologi, Universitas Muhammadiyah Surakarta.
14. Shidiq, A.S., Masykuri, M., dan Susanti V. H., E. 2014. Pengembangan Penilaian Instrumen Two-Tier Multiple Choice Untuk Mengukur Kemampuan Berpikir Tingkat Tinggi (Higher Order Thinking Skills) Pada Materi Kelarutan dan Hasil Kali Kelarutan Untuk Siswa SMA/MA Kelas XI. *Jurnal Pendidikan Kimia*. 3 (1), 83-92.
15. Sugiyono, (2012). *Metode Penelitian Kuantitatif Kualitatif R&D*. Bandung, Alfabeta.
16. Sumintono, B., dan Widhiarso, W. 2015. Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial. Cimahi: Trimkomunikata.
17. Treagust, David F. (2006). Diagnostic Assesment In Science as A Means to Improving Teaching, Learning, and Retention. *Uni Serve Science Assesment Symposium Proceedings*. The University of Sydney.
18. Wagiran. 2007. Inovasi Pembelajaran dalam Penyiapan tenaga Kerja Masa Depan. *Jurnal Pendidikan Teknologi & Kejuruan*. 16 (1), 43-55.