

**PENGEMBANGAN INSTRUMEN TES KETERAMPILAN BERPIKIR
TINGKAT TINGGI PADA MATERI USAHA DAN ENERGI**

(Skripsi)

Oleh

AHMAD NUR FUADI



**PENDIDIKAN FISIKA
FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2019**

ABSTRAK

PENGEMBANGAN INSTRUMEN TES KETERAMPILAN BERPIKIR TINGKAT TINGGI PADA MATERI USAHA DAN ENERGI

Oleh

AHMAD NUR FUADI

Penelitian ini bertujuan untuk menghasilkan instrumen tes keterampilan berpikir tingkat tinggi materi usaha dan energi dalam bentuk *Two-Tier Multiple Choice*. Metode yang digunakan ialah metode pengembangan instrumen tes oleh Istiyono dkk. (2014) yang secara umum terdiri dari tiga tahapan diantaranya tahap perancangan tes, uji coba tes dan perakitan tes. Pada salahsatu tahap perancangan tes dilakukan uji ahli kepada tiga orang ahli yang menguasai materi usaha dan energi serta ahli pembelajaran fisika. Setelah mendapatkan hasil uji dari ahli instrumen tes diujicobakan secara terbatas. Subjek yang digunakan dalam uji coba ini berjumlah 32 siswa. Data yang telah diperoleh, kemudian dianalisis menggunakan Model *Rasch* dengan aplikasi *Ministep Winstep*. Berdasarkan hasil analisis Model *Rasch* dapat disimpulkan bahwa instrumen tes keterampilan berpikir tingkat tinggi yang dikembangkan valid dan realibel dengan nilai reliabilitas sebesar 0,77 yang berarti memiliki realibilitas tinggi, dengan demikian instrumen tes memiliki kelayakan sebagai instrumen penilaian.

Kata Kunci: Keterampilan Berpikir Tingkat Tinggi, Instrumen Tes, Usaha dan Energi, Model *Rasch*

**PENGEMBANGAN INSTRUMEN TES KETERAMPILAN BERPIKIR
TINGKAT TINGGI PADA MATERI USAHA DAN ENERGI**

Oleh

AHMAD NUR FUADI

Skripsi

**Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA PENDIDIKAN**

Pada

**Program Studi Pendidikan Fisika
Jurusan Pendidikan Matematika dan Ilmu Pengetahuan Alam**



**PENDIDIKAN FISIKA
FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2019**

Judul Skripsi : **PENGEMBANGAN INSTRUMEN TES
KETERAMPILAN BERPIKIR TINGKAT
TINGGI PADA MATERI USAHA DAN
ENERGI**

Nama Mahasiswa : Ahmad Nur Fuadi

Nomor Pokok Mahasiswa : 1413022003

Program Studi : Pendidikan Fisika

Jurusan : Pendidikan Matematika dan Ilmu Pengetahuan
Alam

Fakultas : Keguruan dan Ilmu Pendidikan



Drs. Nengah Maharta, M.Si.
NIP 19551213 198303 1 022

Wayan Suana, S.Pd., M.Si.
NIP 19851231 200812 1 001

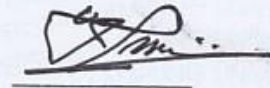
2. Ketua Jurusan Pendidikan MIPA

Dr. Caswita, M.Si.
NIP 19671004 199303 1 004

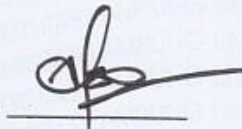
MENGESAHKAN

1. Tim Penguji

Ketua : Drs. Nengah Maharta, M.Si.



Sekretaris : Wayan Suana, S.Pd., M.Si.



Penguji
Bukan Pembimbing : Drs. Feriansyah Sesunan, M.Pd.



2. Dekan Fakultas Keguruan dan Ilmu Pendidikan

Prof. Dr. Patuan Raja, M.Pd. 
NIP. 19620804 198905 1 001

Tanggal Lulus Ujian Skripsi: 25 November 2019

SURAT PERNYATAAN

Yang bertanda tangan di bawah ini adalah:

Nama : Ahmad Nur Fuadi

NPM : 1413022003

Fakultas / Jurusan : KIP / Pendidikan MIPA

Program Studi : Pendidikan Fisika

Alamat : Desa Karangrejo Kec.Negerikaton Kab.Pesawaran

Dengan ini menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebut dalam daftar pustaka.

Bandarlampung, 09 Desember 2019
Yang Menyatakan,



Ahmad Nur Fuadi
NPM. 1413022003

RIWAYAT HIDUP

Penulis dilahirkan di Karangrejo pada tanggal 21 Desember 1996, anak kedua dari tiga bersaudara dari pasangan Bapak Aziz Muslim dan Ibu Alfiah.

Penulis mengawali pendidikan formal pada tahun 2002 di SDN 3 Pujo Rahayu, diselesaikan tahun 2008. Selanjutnya penulis melanjutkan pendidikan di SMP Negeri 2 Negerikaton hingga tahun 2011, kemudian penulis melanjutkan pendidikannya di MA Negeri 1 Pringsewu, diselesaikan pada tahun 2014. Pada tahun yang sama, penulis diterima dan terdaftar sebagai mahasiswa Program Studi Pendidikan Fisika, Jurusan Pendidikan MIPA, Fakultas Keguruan dan Ilmu Pendidikan di Universitas Lampung melalui Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN).

Pada tahun 2017, penulis melakukan Program Pengalaman Lapangan (PPL) di SMP Negeri 2 Blambangan Umpu Kabupaten Way Kanan dan Kuliah Kerja Nyata (KKN) di Kampung Umpu Bhakti Kecamatan Blambangan Umpu Kabupaten Way Kanan.

MOTTO

“Sebaik-baik Manusia Adalah yang Bermanfaat Bagi Orang Lain”

(Rasululloh)

“Menyesali Nasib Tidak akan Mengubah Keadaan. Terus Berkarya dan Bekerjalah yang Membuat Kita Berharga”(KH. Abdurrahman Wahid)

PERSEMBAHAN

Teriring do'a dan rasa syukur kehadiran Allah Subhanahu wata'ala yang selalu melimpahkan rahmat dan karunia-Nya. Ku persembahkan lembaran-lembaran sederhana ini sebagai tanda cinta dan kasihku yang tulus kepada:

1. Bapak Aziz Muslim dan Ibu Alfiah tercinta yang telah tulus berkorban, membimbing, menyayangi, dan mendoakan setiap waktu untuk keberhasilanku dunia dan akhirat.
2. Kakakku tercinta Nikmatul Khoiriyah yang selalu menyayangiku, memperhatikanku, mendukungku, telah memberikan doa dan semangatnya untuk keberhasilanku.
3. Adikku tercinta Saadatul Azizah yang telah memberiku semangat dan keceriaan,
4. Keluarga besarku yang selalu mendukung, mendoakan, dan membantu keberhasilanku.
5. Para pendidik yang telah memberikanku banyak ilmu baik ilmu dunia maupun ilmu agama.
6. Keluarga besar Fighter 14 (Pendidikan Fisika 2014)
7. Almamaterku tercinta Universitas Lampung.

SANWACANA

Bismillahirrohmanirrohim

Puji dan syukur penulis haturkan kepada Allah Subhanahu wata'ala, yang selalu melimpahkan rahmat dan kasih sayang-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul “Pengembangan Instrumen Tes Keterampilan Berpikir Tingkat Tinggi pada Materi Usaha dan Energi” sebagai salah satu syarat untuk memperoleh gelar Sarjana Pendidikan di Universitas Lampung.

Pada kesempatan ini penulis mengucapkan terimakasih kepada:

1. Bapak Prof. Dr. Patuan Raja, M.Pd., selaku Dekan Fakultas Keguruan dan Ilmu Pendidikan Universitas Lampung.
2. Bapak Dr. Caswita, M.Si., selaku Ketua Jurusan Pendidikan MIPA.
3. Bapak Drs. I Wayan Distrik, M.Si., selaku Ketua Program Studi Pendidikan Fisika.
4. Bapak Drs. Nengah Maharta, M.Si., selaku dosen Pembimbing Akademik dan Pembimbing I yang telah memotivasi, membimbing, dan mengarahkan penulis selama penulisan skripsi.
5. Bapak Wayan Suana, S.Pd., M.Si., selaku Pembimbing II, atas kesabarannya dalam memberikan bimbingan dan motivasi kepada penulis.
6. Bapak Drs. Feriansyah Sesunan, M.Pd., selaku Pembahas yang banyak memberikan masukan dan kritik yang bersifat positif dan membangun.

7. Bapak dan ibu dosen Pendidikan Fisika Universitas Lampung yang telah membimbing penulis dalam pembelajaran di Universitas Lampung.
8. Sahabat-sahabatku keluarga besar pendidikan fisika 2014 A dan keluarga besar pendidikan fisika 2014 B.
9. Kepada semua pihak yang telah membantu terselesainya skripsi ini.

Penulis berdoa, semoga semua amal dan bantuan mendapat pahala serta balasan dari Allah Subhanahu Wata'ala dan semoga skripsi ini bermanfaat bagi dunia pendidikan. Aamiin.

Bandar Lampung, 09 Desember 2019

Penulis

DAFTAR ISI

	Halaman
DAFTAR TABEL	vi
I. PENDAHULUAN	
A. Latar Belakang Masalah.....	1
B. Rumusan Masalah	5
C. Tujuan Penelitian	6
D. Manfaat Penelitian	6
E. Ruang Lingkup Penelitian.....	6
II. TINJAUAN PUSTAKA	
A. Kemampuan Berpikir Tingkat Tinggi	8
B. Instrumen Tes	12
C. Prosedur Pengembangan Instrumen Tes	20
III. METODE PENELITIAN	
A. Metode Penelitian.....	27
B. Prosedur Penelitian.....	28
IV. HASIL DAN PEMBAHASAN	
A. Hasil Pengembangan	36
1. Hasil Uji Ahli	38
2. Hasil Uji Keterbacaan	40
3. Hasil Uji Coba Terbatas	42
B. Pembahasan	45

V. SIMPULAN

A. Simpulan	49
B. Saran.....	50
DAFTAR PUSTAKA	51
LAMPIRAN.....	57
Lampiran 1.Kisi-Kisi Instrumen Tes Keterampilan Berpikir Tingkat Tinggi	57
Lampiran 2.Instrumen Tes Keterampilan Berpikir Tingkat Tinggi	66
Lampiran 3.Kunci Jawaban Tes Keterampilan Berpikir Tingkat Tinggi.....	81
Lampiran 4.Instrumen Uji Ahli	103
Lampiran 5.Rekapitulasi Hasil Uji Ahli.....	109
Lampiran 6.Rekapitulasi Skor Hasil Tes Keterampilan Berpikir Tingkat Tinggi.....	110
Lampiran 7.Hasil Analisis Data <i>Rasch</i> Model Uji Coba	111

DAFTAR TABEL

Tabel	Halaman
1. Perbedaan Taksonomi Bloom dan Anderson Keterampilan Berpikir Tingkat Tinggi	10
2. Deskripsi Kata Kunci dan Kata Kerja Operasional Revisi Taksonomi Bloom Untuk Keterampilan Berpikir Tingkat Tinggi	11
3. Pedoman Penskoran	30
4. Kriteria Penilaian untuk Hasil Validasi Ahli dan Uji Coba Produk.....	31
5. Kriteria Hasil Evaluasi Validitas Instrumen	31
6. Nilai <i>Person</i> dan item <i>reability</i>	34
7. Nilai <i>α Cronbach</i>	35
8. Indikator Berpikir Tingkat Tinggi yang Dikembangkan.....	37
9. Hasil Validasi Ahli	39
10. Hasil Uji Keterbacaan Siswa	41
11. Analisis Kesesuaian Butir	43

I. PENDAHULUAN

A. Latar Belakang Masalah

Perkembangan zaman terus berlangsung hingga saat ini dunia berada pada zaman globalisasi. Pada zaman ini persaingan di segala bidang cukup ketat, hal yang mendukung persaingan tersebut adalah kualitas Sumber Daya Manusia (SDM). Kualitas SDM suatu bangsa ditentukan oleh tingkat pendidikan bangsa tersebut (Istiyono dkk., 2014). Peningkatkan kualitas pendidikan dimulai dari peningkatan kualitas pembelajaran. Peningkatkan kualitas pembelajaran dapat dimulai dengan menyusun tujuan pembelajaran yang tepat.

Berdasarkan standar isi Kurikulum 2013 (Depdikbud, 2013: 69), salah satu tujuan pembelajaran Fisika SMA adalah agar peserta didik memiliki keterampilan berpikir kognitif dan berpikir kritis sehingga dapat dikatakan tujuan pembelajaran Fisika SMA di Indonesia adalah agar siswa memiliki keterampilan berpikir tingkat tinggi yang mendukung persaingan di era globalisasi. Oleh sebab itu guru dituntut untuk dapat mendesain pembelajaran yang mampu memfasilitasi siswa untuk memiliki keterampilan berpikir tingkat tinggi. Salah satu tugas yang sangat penting dilakukan guru dalam mendesain pembelajaran adalah mengembangkan instrumen penilaian karena proses pembelajaran yang baik akan tercermin dari penilaian hasil belajar

yang dilakukan dengan menggunakan instrumen penilaian yang tepat dan sesuai dengan tujuan pembelajaran atau kompetensi yang telah ditetapkan.

Keterampilan berpikir tingkat tinggi merupakan keterampilan yang penting dimiliki di zaman globalisasi yang memasuki abad 21 ini sehingga perlu diajarkan kepada siswa. Hal ini sejalan dengan pendapat Greenstein (2012) 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 keterampilan berpikir tingkat tinggi dan pemecahan masalah (Trilling & Fadel, 2009).

Berkenaan dengan keterampilan berpikir tingkat tinggi ini, menurut Rofiah dkk. (2013) berdasarkan hasil salah satu studi internasional mengenai keterampilan kognitif siswa dalam TIMSS (*Trends in Mathematics and Science Study*) yang diadakan oleh IEA (*International Association for the Evaluation of Educational Achievement*) keterampilan berpikir tingkat tinggi siswa Indonesia masih rendah. Hal ini dapat disebabkan karena salahsatunya model asesmen yang tidak tepat. Salahsatu cara memperbaiki kualitas pembelajaran dapat dilakukan dengan memperbaiki evaluasi pembelajaran yang diberikan (Dahlan, 2016). Asesmen yang tepat dapat mendorong siswa untuk belajar dengan berpikir tingkat tinggi (Istiyono dkk., 2014).

Sebagaimana Barnett & Francis dalam Himah dkk. (2016) berpendapat bahwa pertanyaan berpikir tingkat tinggi dapat mendorong siswa untuk

berpikir secara mendalam tentang materi pelajaran, sehingga dapat dikatakan bahwa tes keterampilan berpikir tingkat tinggi dapat memberikan rangsangan kepada siswa untuk mengembangkan keterampilan berpikir tingkat tingginya.

Selain itu sejalan dengan hasil penelitian Istiyono dkk. (2014) sebuah studi pendahuluan yang dilakukan Sugiarti dkk. (2017) menunjukkan bahwa instrumen penilaian fisika di sekolah hanya mengukur keterampilan berpikir tingkat rendah siswa. Keterampilan berpikir kritis, yang mencakup salah satu keterampilan berpikir tingkat tinggi, belum diperiksa oleh para guru. Hal ini dibuktikan dengan hasil temuan yang mengacu pada lembar butir soal ujian nasional. Analisis lembar butir soal ujian nasional dalam lima tahun terakhir, menunjukkan bahwa 7% dari soal tersebut memeriksa memori, 28% mengukur pemahaman, dan 62,4% mengukur penerapan. Ketiga aspek tersebut mencakup keterampilan berpikir tingkat rendah (Sugiarti dkk., 2017).

Penyusunan alat evaluasi yang mengacu pada keterampilan berpikir tingkat tinggi sangat perlu dikembangkan dan diujikan pada siswa terutama di jenjang Sekolah Menengah Atas karena evaluasi yang dilakukan oleh TIMSS dan PISA hanya untuk usia SD dan SMP (Ramadhan dkk., 2018). Namun demikian di lapangan masih banyak dijumpai guru yang belum mampu dalam mengembangkan perangkat yang mendukung pencapaian keterampilan berpikir tingkat tinggi siswa di sekolah. Sampai saat ini guru-guru masih memerlukan pelatihan dan bimbingan dalam mengembangkan perangkat tes yang mendukung kemampuan siswa untuk berpikir tingkat tinggi (Siswoyo & Sunaryo, 2017).

Sejalan dengan keadaan tersebut, nyatanya di MA Negeri 1 Pringsewu belum ada instrumen khusus yang digunakan untuk mengukur keterampilan 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 instrumen tes yang masih tergolong tes keterampilan berpikir tingkat rendah dengan tipe esay dan pilihan ganda.

Instrumen tes untuk mengukur keterampilan berpikir tingkat tinggi dapat menggunakan berbagai bentuk penilaian, salahsatunya *multiple choice*.

Bentuk tes *multiple choice* dapat dimodifikasi untuk mendukung pengukuran keterampilan berpikir tingkat tinggi. Salah satu alternatif bentuk modifikasi *multiple choice* yang dapat digunakan adalah *two-tier multiple choice* (TTMC) test atau pilihan ganda bertingkat. Keunggulan *two-tier multiple choice test* yaitu dapat mengurangi tingkat kecenderungan siswa menebak jawaban dengan mengungkapkan alasan siswa memilih opsi jawaban (Dewati dkk., 2016). Penyertaan alasan pada tingkatan kedua dalam *two-tier multiple choice* dapat digunakan untuk melihat keterampilan siswa dalam memberi alasan dan meningkatkan keterampilan berpikir tingkat tinggi (Cullinane & Maeve, 2011).

Agar instrumen tes yang dikembangkan dapat mengukur keterampilan siswa dengan baik maka perlu diadakan analisis validitas dan realibilitas butir soal. Validitas dan realibiitas butir soal instrumen tes dalam pengembangan ini

dianalisis menggunakan *Model Rasch* dengan bantuan aplikasi *Ministep Winstep*. Pendekatan *Model Rasch* digunakan untuk menganalisis validitas dan reliabilitas instrumen yang dikembangkan (Napitupulu, 2017). Searing dalam Suryani (2018) menjelaskan pengukuran *Rasch model* adalah suatu model pengukuran yang terbentuk berdasarkan hasil pertimbangan validitas dan reliabilitas setiap responden yang menjawab soal dan kesukaran soal bagi setiap soal. Penggunaan model Rasch dalam validasi instrumen akan menghasilkan informasi yang lebih memenuhi definisi pengukuran dan lebih holistik tentang instrumen (Napitupulu, 2017).

Berdasarkan uraian tersebut, untuk mempersiapkan SDM yang mampu bersaing di era globalisasi dan dapat melengkapi tuntutan dalam pembelajaran abad 21 maka telah dilakukan pengembangan instrumen tes *two-tier multiple choice* untuk mengukur keterampilan berpikir tingkat tinggi pada materi Usaha dan Energi Fisika SMA.

B. Rumusan Masalah

Adapun rumusan masalah dalam penelitian ini adalah:

1. Bagaimana validitas instrumen tes keterampilan berpikir tingkat tinggi pada materi usaha dan energi fisika SMA?
2. Bagaimana realibilitas instrumen tes keterampilan berpikir tingkat tinggi pada materi usaha dan energi fisika SMA?

C. Tujuan Penelitian

Adapun Tujuan dari penelitian ini adalah mengembangkan instrumen tes keterampilan berpikir tingkat tinggi sebagai instrumen alternatif untuk mengukur keterampilan berpikir tingkat tinggi pada materi usaha dan energi dalam fisika SMA, serta:

1. Mengetahui validitas instrumen tes keterampilan berpikir tingkat tinggi pada materi usaha dan energi fisika SMA.
2. Mengetahui realibilitas instrumen tes keterampilan berpikir tingkat tinggi pada materi usaha dan energi fisika SMA.

D. Manfaat Penelitian

Penelitian ini memberikan bermanfaat untuk membantu pendidik dalam melakukan penilaian keterampilan berpikir tingkat tinggi siswa dengan tersedianya instrumen tes keterampilan berpikir tingkat tinggi pada materi usaha dan energi fisika SMA dan dapat dijadikan rujukan untuk penelitian-penelitian selanjutnya.

E. Ruang Lingkup Penelitian

Adapun ruang lingkup penelitian pengembangan ini adalah sebagai berikut:

1. Pengembangan yang dimaksud dalam penelitian adalah mengembangkan instrumen tes keterampilan berpikir tingkat tinggi tercetak yang berfungsi

untuk mengukur keterampilan berpikir tingkat tinggi pada materi usaha dan energi fisika SMA.

2. Instrumen tes keterampilan berpikir tingkat tinggi yang dibuat mengindikasikan level berpikir tingkat tinggi berdasarkan klasifikasi keterampilan berpikir tingkat tinggi menurut taksonomi bloom yang telah direvisi.
3. Validasi produk yang telah dikembangkan dilakukan dengan uji ahli, yang meliputi aspek materi, desain dan bahasa.
4. Uji lapangan untuk melihat instrumen tes keterampilan berpikir tingkat tinggi yang baik digunakan dalam penilaian dilakukan dengan mengimplementasikan produk dalam pelajaran fisika di MAN 1 Pringsewu tahun ajaran 2018/2019 kelas X MIA.
5. Analisis data yang digunakan menggunakan analisis data *Rasch Model* dengan bantuan aplikasi *Ministep Winstep*.

DAFTAR ISI

	Halaman
DAFTAR TABEL	vi
I. PENDAHULUAN	
A. Latar Belakang Masalah.....	1
B. Rumusan Masalah	5
C. Tujuan Penelitian	5
D. Manfaat Penelitian	6
E. Ruang Lingkup Penelitian.....	6
II. TINJAUAN PUSTAKA	
A. Kemampuan Berpikir Tingkat Tinggi	8
B. Instrumen Tes	12
C. Prosedur Pengembangan Instrumen Tes	20
III. METODE PENELITIAN	
A. Metode Penelitian.....	27
B. Prosedur Penelitian.....	28
IV. HASIL DAN PEMBAHASAN	
A. Hasil Pengembangan	36
1. Hasil Uji Ahli	38
2. Hasil Uji Keterbacaan	40
3. Hasil Uji Coba Terbatas	42
B. Pembahasan.....	45

V. SIMPULAN

A. Simpulan	49
B. Saran.....	50
DAFTAR PUSTAKA	51
LAMPIRAN	57
Lampiran 1. Kisi-Kisi Instrumen Tes Keterampilan Berpikir Tingkat Tinggi	57
Lampiran 2. Instrumen Tes Keterampilan Berpikir Tingkat Tinggi	66
Lampiran 3. Kunci Jawaban Tes Keterampilan Berpikir Tingkat Tinggi.....	81
Lampiran 4. Instrumen Uji Ahli	103
Lampiran 5. Rekapitulasi Hasil Uji Ahli.....	109
Lampiran 6. Rekapitulasi Skor Hasil Tes Keterampilan Berpikir Tingkat Tinggi.....	110
Lampiran 7. Hasil Analisis Data <i>Rasch</i> Model Uji Coba	111

DAFTAR TABEL

Tabel	Halaman
1. Perbedaan Taksonomi Bloom dan Anderson Keterampilan Berpikir Tingkat Tinggi	10
2. Deskripsi Kata Kunci dan Kata Kerja Operasional Revisi Taksonomi Bloom Untuk Keterampilan Berpikir Tingkat Tinggi	11
3. Pedoman Penskoran	30
4. Kriteria Penilaian untuk Hasil Validasi Ahli dan Uji Coba Produk.....	31
5. Kriteria Hasil Evaluasi Validitas Instrumen	31
6. Nilai <i>Person</i> dan item <i>reability</i>	34
7. Nilai <i>Cronbach</i>	35
8. Indikator Berpikir Tingkat Tinggi yang Dikembangkan.....	37
9. Hasil Validasi Ahli	39
10. Hasil Uji Keterbacaan Siswa.....	41
11. Analisis Kesesuaian Butir	44

II. TINJAUAN PUSTAKA

A. Keterampilan Berpikir Tingkat Tinggi

Kemampuan berpikir tingkat tinggi yang dalam bahasa Inggrisnya *Higher order Thinking Skill* (HOTS) menurut Siswyo dan Sunaryo (2017) adalah konsep reformasi pendidikan berdasarkan Taksonomi Bloom. Idennya adalah bahwa beberapa jenis pembelajaran memerlukan pengolahan lebih kognitif daripada yang lain, tetapi juga memiliki manfaat yang lebih umum. Menurut Iskandar dan Senam (2015) *Higher order thinking skills* (HOTS) atau keterampilan berpikir tingkat tinggi merupakan bagian dari taksonomi Bloom hasil revisi yang berupa kata kerja operasional yang terdiri dari *analyze* (C4), *evaluate* (C5) dan *create* (C6) yang dapat digunakan dalam penyusunan soal. Menurut Anderson dan Karthwholl dalam Himah dkk. (2016) taksonomi Bloom membagi proses kognitif menjadi keterampilan berpikir tingkat rendah atau *Lower Order Thinking* (LOT) dan keterampilan berpikir tingkat tinggi atau *Higher Order Thinking* (HOT). Keterampilan yang termasuk LOT adalah keterampilan mengingat (*remember*), memahami (*understand*), dan menerapkan (*apply*), sedangkan HOT meliputi keterampilan menganalisis (*analyze*), mengevaluasi (*evaluate*), dan menciptakan (*create*).

Menurut Malik dkk. (2015) HOTS adalah keterampilan berpikir untuk memeriksa, menghubungkan, dan mengevaluasi semua aspek situasi dan

masalah. Termasuk di dalamnya mengumpulkan, mengorganisir, mengingat, dan menganalisa informasi. Keterampilan berpikir tingkat tinggi merupakan keterampilan menghubungkan, memanipulasi, dan mentransformasi pengetahuan serta pengalaman yang sudah dimiliki untuk berpikir secara kritis dan kreatif dalam upaya menentukan keputusan dan memecahkan masalah pada situasi baru (Rofiah dkk., 2013).

Selanjutnya Rofiah dkk.(2013) mengemukakan Secara umum, terdapat beberapa aspek yang menunjukkan keterampilan berpikir tingkat tinggi yang dimiliki oleh seseorang yaitu keterampilan berpikir kritis, berpikir kreatif, serta memecahkan masalah. Johnson (2007) mengemukakan bahwa berpikir kritis adalah sebuah proses terorganisasi yang memungkinkan siswa mengevaluasi bukti, asumsi, logika, dan bahasa yang mendasari pemikiran orang lain. Begitu juga menurut Sari dkk. (2016) Berpikir kritis merupakan salah satu ketrampilan berpikir tingkat tinggi (*Higher Order Thinking Skill*), dengan skala proses lebih luas yang dapat merefleksikan perilaku manusia. Menurut Heong dkk. (2011) keterampilan berpikir tingkat tinggi ini menghendaki seseorang untuk menerapkan informasi baru atau pengetahuan sebelumnya dan memanipulasi untuk menjangkau kemungkinan jawaban dalam situasi baru.

Berdasarkan beberapa pendapat di atas dapat ditarik suatu pengertian bahwasannya keterampilan berpikir tingkat tinggi adalah keterampilan berpikir dengan ilmu yang telah diperoleh untuk menganalisis, mengevaluasi

dan mensintesis suatu permasalahan baru yang dihadapi sehingga permasalahan baru tersebut mampu terpecahkan.

Dalam Taksonomi Bloom yang telah direvisi keterampilan berpikir yang melibatkan analisis (C4), mengevaluasi (C5), dan mencipta atau kreativitas (C6) dianggap berpikir tingkat tinggi (Krathworl & Anderson, 2001). Anderson telah melakukan penelitian serta didapatkan perbaikan dalam Taksonomi Bloom yang sudah ada. Perbaikan tersebut, yaitu mengubah Taksonomi Bloom dari kata benda menjadi kata kerja. Revisi taksonomi Bloom ini menekankan pada kata kerja, mengubah kata “*knowledge*” pada Taksonomi Bloom lama menjadi “*remember*”, karena “*knowledge*” dipandang sebagai kata benda yang maknanya lebih luas. Selanjutnya sebagai kata benda, *knowledge* dikategorikan sebagai dimensi yang terpisah. Hal ini dilakukan karena Taksonomi Bloom yang sebenarnya yaitu penggambaran proses berpikir. Setelah itu dilakukanlah pergeseran susunan taksonomi bloom yang menjabarkan berpikir tingkat rendah dan berpikir tingkat tinggi. Perbedaan Taksonomi Bloom dan Anderson Kemampuan Berpikir Tingkat Tinggi dapat dilihat pada tabel 1.

Tabel 1 Perbedaan Taksonomi Bloom dan Anderson Keterampilan Berpikir Tingkat Tinggi

Taksonomi Bloom	Revisi Taksonomi Bloom	Keterangan
Analisis	Menganalisis	C4
Sintesis	Menilai	C5
Penilaian	Menciptakan	C6

(Anderson & Krathworl, 2001)

Menurut Anderson dan Krathworl dalam Yusmanto dkk. (2016) Deskripsi kata kunci dan kata kerja operasional setiap kategori revisi taksonomi Bloom dan Anderson dapat dilihat dalam Tabel 2.

Tabel 2. Deskripsi Kata Kunci dan Kata Kerja Operasional Revisi Taksonomi Bloom untuk Keterampilan Berpikir Tingkat Tinggi

Deskripsi Kata Kunci Revisi Taksonomi Bloom	Kata Kerja Operasional Revisi Taksonomi Bloom
Menganalisis: dapatkah peserta didik membedakan antara konsep-konsep yang berbeda?	Menilai, membandingkan, mengurutkan, membedakan, menentukan.
Mengevaluasi: dapatkah peserta didik membenarkan suatu pernyataan atau pilihan tertentu dengan memberikan alasan?	Mengevaluasi, menilai, mengkritik, memilih/menyeleksi, menghubungkan, memberikan pendapat.
Mencipta: dapatkah peserta didik membuat atau mengembangkan produk, teori atau sudut pandang baru berdasarkan pembelajaran?	Merakit, mendesain, merancang, membuat, memformulasikan, memberikan solusi.

(Yusmanto dkk., 2016)

Dalam taksonomi Bloom domain kognitif hanya terdiri dari satu dimensi saja namun dalam taksonomi Anderson dan Krathwohl berubah menjadi dua dimensi. Dimensi yang pertama adalah *Knowledge Dimension* (dimensi pengetahuan) dan *Cognitive Process Dimension* (dimensi proses kognisi). Dimensi proses kognisi terdapat 6 kategori, yaitu keterampilan mengingat, memahami, dan menerapkan yang merupakan keterampilan berpikir tingkat rendah. Selain itu keterampilan menganalisis, mengevaluasi, dan mencipta termasuk dalam keterampilan berpikir tingkat tinggi. Menurut Krathwohl dalam Lewy dkk. (2009) menjelaskan bahwa indikator untuk mengukur keterampilan berpikir tingkat tinggi meliputi :

a. Menganalisis

- 1) Menganalisis informasi yang masuk dan membagi-bagi atau menstrukturkan informasi ke dalam bagian yang lebih kecil untuk mengenali pola atau hubungannya.
- 2) Mampu mengenali serta membedakan faktor penyebab dan akibat dari sebuah fenomena yang rumit.
- 3) Mengidentifikasi atau merumuskan pertanyaan

b. Mengevaluasi

- 1) Memberikan penilaian terhadap solusi, gagasan, dan metodologi dengan menggunakan kriteria yang cocok atau standar yang ada untuk memastikan nilai efektivitas atau manfaatnya.
- 2) Membuat hipotesis, mengkritik dan melakukan pengujian.
- 3) Menerima atau menolak suatu pernyataan berdasarkan kriteria yang telah ditetapkan

c. Mengkreasi

- 1) Membuat generalisasi suatu ide atau cara pandang terhadap sesuatu.
- 2) Merancang suatu cara untuk menyelesaikan masalah.
- 3) Mengorganisasikan unsur-unsur atau bagian-bagian menjadi struktur baru yang belum pernah ada sebelumnya.

B. Instrumen Tes

Guru perlu memantau proses, kemajuan, dan perbaikan hasil belajar peserta didik secara berkesinambungan dengan menggunakan penilaian. Penilaian pendidikan adalah proses pengumpulan dan pengolahan informasi untuk

menentukan pencapaian hasil belajar peserta didik (Depdiknas, 2007: 2).

Tujuan tes yang paling penting untuk mengetahui tingkat kemampuan siswa, selain itu juga dapat digunakan untuk mengasah kemampuan berpikir tingkat tinggi siswa (Ekawati dkk., 2017). Kemendikbud (2017) menjelaskan bahwa instrumen tes keterampilan berpikir tingkat tinggi memiliki peran dalam penilaian untuk mempersiapkan kompetensi peserta didik menyongsong abad 21, memupuk rasa cinta dan peduli terhadap kemajuan daerah, meningkatkan motivasi belajar peserta didik dan meningkatkan mutu pendidikan.

Penilaian pendidikan dapat dilakukan secara lisan ataupun tertulis. Penilaian secara tertulis dilakukan dengan tes tertulis. Secara garis besar ada dua bentuk soal tes tertulis, yaitu: memilih jawaban dan mensuplai jawaban. Soal tes tertulis yang menjawabnya dengan memilih jawaban antara lain: pilihan ganda, dua pilihan (benar-salah, ya-tidak), menjodohkan, dan sebab-akibat.

Istiyono dkk. (2014) mengungkapkan kenyataan di lapangan bahwa tes pilihan ganda lebih banyak digunakan dari pada bentuk tes yang lain. Hal ini sejalan dengan Saputro dan Supahar (2018) yang menyatakan tes tertulis berbentuk pilihan ganda merupakan bentuk tes yang paling umum digunakan untuk mengukur kemampuan berpikir pada mata pelajaran fisika peserta didik SMA di Indonesia. Tes pilihan ganda adalah bentuk tes obyektif yang mempunyai ciri utama kunci jawaban jelas dan pasti sehingga hasilnya dapat diskor secara obyektif. Artinya setelah siswa mengerjakan soal dalam bentuk tes pilihan ganda maka siswa tersebut akan memperoleh skor yang sama jika hasil pekerjaannya diperiksa oleh lebih dari satu pemeriksa. Hal ini disebabkan

setiap jawaban diberi skor yang sudah pasti dan tidak mengenal jawaban di antara benar dan salah atau jawaban benar sebagian saja. Soal pilihan ganda terdiri dari pernyataan dan pertanyaan yang harus dijawab oleh siswa atau melengkapi dengan memilih salah satu dari beberapa alternatif yang tersedia. Satu di antaranya adalah yang paling benar, lainnya disebut pengecoh (Kadir, A. 2015). Tes objektif pilihan ganda memiliki beberapa kebaikan yang tidak dimiliki tes jenis lainnya, diantaranya adalah bentuk tes pilhan ganda lebih representatif dalam hal mencakup dan mewakili materi pembelajaran., lebih memungkinkan bagi tester untuk bertindak lebih objektif, baik dalam, mengoreksi lembar jawaban soal, menentukan bobot skor, maupun dalam menentukan nilai akhirnya, butir-butir soal jauh lebih mudah dianalisis, baik analisis dari segi tingkat kesukaran, daya pembeda, maupun reliabilitasnya,dan dari segi efisiensi waktu, mengoreksi tes objektif pilihan ganda jauh lebih cepat (Ivanty dkk., 2013)

Selanjutnya Istiyono dkk. (2014) mengungkapkan berdasarkan hasil survei pendahuluan dengan cara melakukan wawancara dengan guru-guru fisika SMA di Daerah Istimewa Yogyakarta (DIY), diperoleh informasi bahwa sebagian besar di sekolah, baik pada tes tengah semester maupun tes akhir semester umumnya menggunakan tes pilihan ganda biasa. Jadi, tes pilihan ganda masih merupakan primadona dalam mendapatkan data prestasi belajar fisika siswa SMA. Pada penelitian pengembangan instrumen tes ini juga akan dibuat tes dengan jenis pilihan ganda yang dimodifikasi yaitu berupa pilihan ganda bertingkat (*two-tier multiple choice*). Penyertaan alasan pada tingkatan kedua soal dapat digunakan untuk mengurangi terjadinya untung-untungan yang

sering menjadi kelemahan dari bentuk soal pilihan ganda biasa. Penilaian soal yang objektif, mudah, dan cepat menjadi keunggulan *two-tier multiple choice test* dibandingkan dengan soal keterampilan berpikir tingkat tinggi yang lainnya contohnya soal essay. Adapun kelemahan dari soal *two-tier multiple choice test* yaitu tidak mampu digunakan untuk mengukur kemampuan verbal siswa seperti esai (Noviana dkk., 2014).

Menurut Arifin dkk. (2017) beberapa hal yang harus diperhatikan, agar instrumen yang dibuat bisa dilihat kriterianya, apakah instrumennya baik atau tidak baik, yaitu validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan distraktor/pengecoh.

1. Validitas

Validitas adalah ketepatan interpretasi yang diperoleh dari hasil penilaian. Validitas berasal dari kata *validity* yang memiliki arti sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Secara khusus validitas dari suatu perangkat tes dapat diartikan merupakan keterampilan suatu tes untuk mengukur apa yang seharusnya diukur (Arifin dkk., 2017).

Ada tiga tipe validitas, yaitu validitas isi, validitas konstruk dan validitas kriteria. Selain itu, terdapat dua macam validitas isi, yaitu validitas kenampakan dan validitas logika. Validitas isi berarti sejauh mana suatu perangkat tes mencerminkan keseluruhan keterampilan yang hendak diukur yang berupa analisis rasional terhadap domain yang hendak diukur. Validitas kenampakan didasarkan pada pertanyaan apakah suatu

butir-butir dalam perangkat tes mengukur aspek yang relevan dengan domainnya. Validitas logika berkaitan dengan keseksamaan batasan pada domain yang hendak diukur, dan merupakan jawaban apakah keseluruhan butir merupakan sampel representatif dari keseluruhan butir yang mungkin dibuat (Arifin dkk., 2017).

Selain validitas produk, terdapat juga validitas butir soal. Validitas butir soal dilakukan dengan analisis model *rasch* menggunakan aplikasi *ministep winstep*. Riswanda (2018) dalam penelitian pengembangannya untuk menentukan validitas butir soal dilihat dari kesesuaian soal (*item fit*) menggunakan parameter nilai outfit dengan ketentuan nilai *Outfit Mean Square* (MNSQ) yang diterima berkisar antara 0,5 sampai dengan 1,5, nilai *Outfit Z-Standard* (ZSTD) yang diterima berkisar antara -2,0 sampai dengan +2,0 dan Nilai *Pt Measure Corr* yang diterima berkisar 0,32 sampai dengan 0,80. Jika butir soal memperoleh nilai yang masuk rentang nilai pada ketiga kriteria tersebut maka dikatakan valid. Sedangkan jika pada ketiga kriteria tersebut tidak terpenuhi, dapat dipastikan bahwa butir soalnya kurang bagus sehingga perlu diperbaiki atau diganti.

2. Realibilitas

Realibilitas adalah kestabilan atau keajegan. Reliabilitas mengacu pada kekonsistenan atau kestabilan hasil penilaian. Jika tesnya memiliki konsistensi yang tinggi, maka tes tersebut akurat, *reproducible*, dan *generalizable* terhadap kesempatan testing dan instrumen yang sama. Sebagai contohnya dapat diilustrasikan dengan seseorang yang diukur

tinggi badannya akan diperoleh hasil yang tidak berubah walaupun menggunakan alat pengukur yang berbeda dan skala yang berbeda. Kaitannya dengan dunia pendidikan, prestasi atau keterampilan seorang siswa dikatakan reliabel jika sudah dilakukan pengukuran. Kereliabelan ini bermakna hasil pengukuran akan sama informasinya, walaupun penguji berbeda, korektornya berbeda atau butir soal yang berbeda tetapi memiliki karakteristik yang sama (Arifin dkk., 2017).

Realibilitas instrumen didapat menggunakan pendekatan model Rasch melalui program *Ministeps Winsteps*. Hasil analisis reliabilitas dapat dilihat pada program *Ministep Winsteps* dengan menggunakan output tabel *summary statistic*. Realibilitas yang ditampilkan adalah realibilitas soal (*item reliability*) dan realibilitas respon (*person reliability*). Realibilitas instrumen secara keseluruhan dapat ditentukan melalui nilai *item reliability* (Widyaningsih dan Yusuf, 2018).

3. Tingkat Kesukaran

Tingkat kesukaran butir soal adalah persentase atau proporsi dari peserta tes untuk menjawab benar suatu butir soal. Besarnya tingkat kesukaran berkisar 0,00 - 1,00. Semakin besar tingkat kesukaran yang diperoleh dari hasil hitungan, berarti semakin mudah soal itu dan soal itu harus direvisi (Arifin dkk., 2017).

4. Daya Pembeda

Daya pembeda soal adalah keterampilan suatu butir soal dapat membedakan siswa yang telah menguasai materi yang ditanyakan dan siswa yang tidak/kurang/belum menguasai materi yang ditanyakan.

Manfaat daya pembeda butir soal adalah seperti berikut ini:

a) Untuk meningkatkan mutu setiap butir soal melalui data empiriknya.

Berdasarkan indeks daya pembeda, setiap butir soal dapat diketahui apakah butir soal itu baik, direvisi, atau ditolak.

b) Untuk mengetahui seberapa jauh setiap butir soal dapat mendeteksi/membedakan keterampilan siswa, yaitu siswa yang telah memahami atau belum memahami materi yang diajarkan guru.

Apabila suatu butir soal tidak dapat membedakan kedua keterampilan siswa itu, maka butir soal itu dimungkinkan seperti berikut ini:

- (1) Kunci jawaban butir soal itu tidak tepat,
- (2) Butir soal itu memiliki 2 atau lebih kunci jawaban yang benar,
- (3) Kompetensi yang diukur tidak jelas,
- (4) Pengecoh tidak berfungsi,
- (5) Materi yang ditanyakan terlalu sulit, sehingga banyak siswa yang menebak,
- (6) Sebagian besar siswa yang memahami materi yang ditanyakan berpikir ada yang salah informasi dalam butir soalnya (Arifin dkk., 2017).

5. Analisis Pengecoh/distraktor

Tes bentuk pilihan ganda terstruktur atas item permasalahan yang ditanyakan (atau pokok soal) dan *option* (atau sejumlah kemungkinan jawaban). *Option* sendiri dibagi menjadi dua, yaitu kunci jawaban dan pengecoh. Ini berarti, dari sekian jawaban yang disIstiyonoakan hanya terdapat satu jawaban yang tepat yang disebut kunci jawaban, sedangkan kemungkinan jawaban yang lain (jawaban yang tidak benar) disebut pengecoh/distraktor. Seseorang yang membuat soal pilihan ganda, terkadang ia tidak mudah untuk membuat pengecohnya. Kecilnya angka daya pembeda seringkali muncul karena pengecoh yang kurang berfungsi. Distraktor atau pengecoh berfungsi sebagai pengidentifikasi peserta tes yang berketerampilan tinggi. Pengecoh dikatakan berfungsi efektif jika banyak dipilih oleh peserta tes dari kelompok bawah, sebaliknya jika banyak dipilih oleh kelompok atas, maka distraktor/pengecoh tersebut tidak berfungsi sebagaimana mestinya (Arifin dkk., 2017).

Dari lima ciri-ciri (karakteristik) tes yang baik di atas, setidaknya ada dua karakteristik yang harus menjadi perhatian dan dianggap paling penting, sehingga sering kali dijadikan dasar dalam menentukan keterpercayaan suatu tes sebagai instrumen, baik sebagai instrumen keberhasilan proses belajar mengajar ataupun sebagai instrument suatu penelitian khususnya untuk data kuantitatif. Kedua karakteristik tersebut adalah validitas (kesahihan) dan reliabilitas (kejegan). Jadi, tes harus *valid* (sahih) dan *reliable* (ajeg) sehingga hasilnya dapat dipertanggungjawabkan. Hal ini sebagaimana Kadir

(2015) dan Novinda (2019) dalam penelitiannya yang menyatakan soal yang memenuhi persyaratan adalah soal yang valid dan reliabel.

C. Prosedur Pengembangan Instrumen Tes

Penelitian dan pengembangan atau yang dikenal dengan istilah *Research and Development* (R&D) dalam bidang pendidikan merupakan model penelitian yang banyak digunakan dalam penelitian pengembangan pendidikan. Sugiyono (2010: 297) menyatakan bahwa metode penelitian dan pengembangan merupakan metode penelitian yang digunakan untuk menghasilkan produk tertentu, dan menguji keefektifan produk tersebut.

Putra dalam Sudaryono dkk. (2013:11) menyatakan bahwa:

Metode penelitian dan pengembangan atau dalam istilah bahasa Inggrisnya *Research and Development* adalah metode penelitian yang digunakan untuk menghasilkan produk tertentu, dan menguji keefektifan produk tersebut. Untuk dapat menghasilkan produk tertentu digunakan penelitian yang bersifat analisis kebutuhan dan menguji keefektifan produk tersebut.

Berdasarkan penjelasan di atas, dapat disimpulkan bahwa metode penelitian dan pengembangan merupakan metode penelitian yang digunakan untuk membuat atau menghasilkan, mengembangkan dan memvalidasi produk tertentu yang bukan untuk menguji suatu teori, kemudian produk tersebut divalidasi dan diuji keefektifannya.

Djaali (2004) menjelaskan mengenai langkah-langkah konstruksi tes dapat ditempuh sebagai berikut:

1. Menetapkan tujuan tes.

Untuk mengembangkan tes sebagai alat evaluasi, hal pertama yang harus dilakukan adalah menetapkan tujuan tes. Setiap tes yang dibuat harus memiliki tujuan yang jelas diantaranya tes untuk klasifikasi kelompok, seleksi ataupun untuk mendiagnosis kesulitan hasil belajar siswa dan lain sebagainya.

2. Analisis kurikulum.

Analisis kurikulum memiliki tujuan untuk menentukan setiap bobot pokok bahasan yang akan dijadikan dasar dalam menentukan jumlah butir soal untuk setiap pokok bahasan soal objektif atau bobot soal yang berbentuk uraian, dalam membuat kisi-kisi tes.

3. Analisis buku pelajaran dan sumber materi belajar lainnya.

Analisis ini memiliki tujuan yang sama dengan analisis kurikulum, yaitu menentukan bobot setiap pokok bahasan. Namun dalam sebuah buku pelajaran yang dianalisis menentukan bobot setiap pokok bahasan sesuai jumlah halaman materi yang terdapat dalam buku pelajaran atau sumber materi belajar lainnya.

4. Membuat kisi-kisi.

Manfaat dalam membuat kisi-kisi adalah untuk menjamin sampel soal yang baik, dalam arti mencakup semua pokok bahasan secara proporsional. Agar butir-butir tes yang dibuat dapat mencakup keseluruhan pokok bahasan atau sub pokok bahasan secara proporsional, maka sebelum menulis butir-butir tes tersebut terlebih dahulu dibuat kisi-kisi sebagai pedoman untuk memudahkan membuat butir-butir tes.

5. Penulisan tujuan instruksional khusus.

Penulisan tujuan instruksional khusus harus sesuai dengan ketentuan yang telah ditetapkan. Tujuan instruksional khusus harus mencerminkan tingkah laku peserta didik, oleh karena itu harus dirumuskan secara operasional, dan secara teknis menggunakan kata-kata operasional (KKO).

6. Penulisan soal

Setelah kisi-kisi dalam bentuk tabel-tabel spesifikasi telah dibuat, maka langkah selanjutnya adalah membuat butir-butir soal. Hal-hal yang harus diperhatikan dalam membuat butir-butir soal adalah sebagai berikut:

- a. Soal yang dibuat harus valid dalam arti mampu mengukur apa yang hendak diukur, yaitu tercapai atau tidaknya tujuan pembelajaran yang telah dirumuskan.
- b. Soal yang dibuat harus dapat dikerjakan dengan menggunakan satu keterampilan spesifik, tanpa dipengaruhi oleh keterampilan lain yang tidak relevan.
- c. Soal yang dibuat harus terlebih dahulu dikerjakan atau diselesaikan dengan langkah-langkah lengkap sebelum digunakan pada tes yang sesungguhnya.
- d. Menetapkan sejak awal aspek keterampilan yang hendak diukur untuk setiap soal yang hendak dibuat
- e. Dalam membuat soal, menghindari sekecil apapun kesalahan dalam pengetikan soal, karena hal tersebut akan mempengaruhi validitas soal.
- f. Memberikan petunjuk mengerjakan soal secara lengkap dan jelas untuk setiap bentuk soal dalam suatu tes.

7. Telaah soal (*face validity*). Soal-soal yang dibuat masih mungkin terjadi kekurangan atau kekeliruan yang menyangkut aspek keterampilan spesifik yang diukur, bahasa yang digunakan, kesalahan dalam pengetikan dan lain sebagainya. Untuk itu, sebelum diperbanyak, maka soal terlebih dahulu harus ditelaah oleh ahli yang memahami materi tes maupun teknik penulisan soal untuk meneliti validitas permukaan dari soal yang dibuat.

8. Reproduksi soal tes terbatas.

Tes yang sudah jadi dan siap untuk diujicobakan, selanjutnya diperbanyak sesuai dengan jumlah kebutuhan, yakni menurut jumlah sampel yang akan diujicobakan atau jumlah peserta yang akan mengerjakan tes tersebut dalam suatu kegiatan uji-coba tes.

9. Uji coba tes

Tes yang sudah diperbanyak akan diuji cobakan pada sejumlah sampel yang telah ditentukan. Sampel uji coba harus mempunyai karakteristik yang kurang lebih sama dengan karakteristik peserta tes yang sesungguhnya.

10. Analisis hasil uji coba

Berdasarkan data hasil uji coba yang dilakukan di lapangan terkait analisis, terutama analisis butir soal yang meliputi validitas butir dan realibilitas diadakan seleksi dengan menggunakan kriteria validitas dan realibilitas tertentu.

11. Revisi soal

Soal-soal yang valid berdasarkan kriteria validitas empirik dikonfirmasi dengan kisi-kisi. Apabila soal-soal yang dibuat sudah

memenuhi syarat dan telah mewakili semua materi yang akan diujikan kepada peserta tes, soal-soal tersebut selanjutnya dirakit menjadi sebuah tes, tetapi apabila soal-soal yang valid belum memenuhi syarat berdasarkan hasil konfirmasi dari kisi-kisi soal, maka dapat dilakukan perbaikan terhadap soal yang diperlukan.

12. Merakit soal menjadi tes

Setelah dikonfirmasi dengan kisi-kisi maka langkah terakhir adalah merakit tes yang terpilih.

Menurut Matondang (2009) ada lima langkah yang dilakukan dalam menyusun instrumen, yaitu: (1) mengidentifikasi komponen program dan indikatornya; (2) membuat kisi kisi yang di dalamnya terdapat indikator, metode pengumpulan data, dan sumber data; (3) menyusun butir instrumen; (4) menyusun kriteria penilaian; (5) menyusun pedoman pengerjaan instrumen.

Selanjutnya, Istiyono dkk. (2014) memaparkan bahwa prosedur pengembangan instrumen berupa tes keterampilan berpikir tingkat tinggi terdiri atas: (1) perancangan tes, langkahnya meliputi penentuan tujuan tes, penentuan kompetensi yang akan diujikan, penentuan materi yang akan diujikan, penyusunan kisi-kisi tes, penulisan item berdasarkan prinsip-prinsip pengembangan tes berpikir tingkat tinggi, validasi item tes, perbaikan item dan perakitan tes, dan penyusunan pedoman penskoran; (2) uji coba tes, langkahnya meliputi penentuan subjek uji coba, pelaksanaan uji coba tes, analisis data hasil uji coba tes; (3) perakitan tes untuk pengukuran.

Selain itu, berdasarkan Kemendikbud (2017) untuk menulis butir soal keterampilan berpikir tingkat tinggi terdapat beberapa langkah diantaranya, yaitu:

1. Menganalisis KD yang dapat dibuat soal-soal keterampilan berpikir tingkat tinggi
Sebelum membuat soal-soal keterampilan berpikir tinggi terlebih dahulu dipilih KD yang dapat dibuat soal-soal HOTS karena tidak semua KD dapat dibuatkan model-model soal HOTS.
2. Menyusun kisi-kisi soal
Kisi-kisi penulisan soal-soal HOTS bertujuan untuk membantu penulisan butir soal HOTS. Secara umum kisi-kisi tersebut diperlukan untuk memandu dalam memilih KD yang dapat dibuat soal-soal HOTS, memilih materi pokok yang terkait KD yang dipilih, merumuskan indikator soal dan menentukan level kognitif.
3. Memilih stimulus yang menarik dan kontekstual
Stimulus yang digunakan dalam penyusunan soal hendaknya menarik dan kontekstual, artinya mendorong peserta didik untuk membaca stimulus. Stimulus yang menarik umumnya baru, belum pernah dibaca peserta didik. Sedangkan stimulus kontekstual berarti stimulus yang sesuai dengan kenyataan dalam kehidupan sehari-hari. Dalam konteks ujian sekolah dapat dipilih stimulus dari lingkungan sekolah atau daerah setempat.
4. Menulis butir pernyataan sesuai dengan kisi-kisi soal
Butir-butir pertanyaan ditulis sesuai dengan kisi-kisi yang dibuat dan stimulus yang dipilih. Dalam penulisan butir ini pada aspek bahasa relatif sama dengan butir soal pada umumnya.

5. Membuat pedoman penskoran atau kunci jawaban

Setiap butir soal HOTS yang ditulis hendaknya dilengkapi dengan pedoman penskoran atau kunci jawaban.

Berdasarkan uraian di atas, maka dapat disimpulkan bahwa dalam penelitian pengembangan yang bertujuan untuk menghasilkan suatu produk harus melalui beberapa tahapan agar produk yang dihasilkan berkualitas baik, bermanfaat, dan dapat digunakan dalam kegiatan pembelajaran, khususnya pembelajaran fisika.

III. METODE PENELITIAN

A. Metode Penelitian

Penelitian dan pengembangan atau *Research and Development (R&D)* adalah metode penelitian yang digunakan untuk menghasilkan produk tertentu dan menguji keefektifan produk tertentu (Sudaryono dkk., 2013). Produk yang akan dikembangkan dan diuji keefektifannya di dalam penelitian ini adalah instrumen penilaian berupa tes untuk mengukur keterampilan berpikir tingkat tinggi siswa dalam bentuk *two-tier multiple choice*. Langkah-langkah pengembangan instrumen menggunakan langkah pengembangan instrumen oleh Istiyono dkk. (2014) yang secara umum terdiri atas perancangan tes, uji coba tes, dan perakitan tes untuk pengukuran.

Tahap perancangan tes meliputi: (1) penentuan tujuan tes, (2) penentuan kompetensi yang diujikan, (3) penentuan materi yang diujikan, (4) penyusunan kisi-kisi tes, (5) penulisan item berdasarkan prinsip-prinsip pengembangan Tes HOT, (6) validasi item tes, (7) perbaikan item dan perakitan tes, dan (8) penyusunan pedoman penskoran. Adapun tahap uji coba tes meliputi: (1) penetapan subjek uji coba (SMA), (2) pelaksanaan uji coba, dan (3) analisis data hasil uji coba. Tahap terakhir dalam pengembangan tes ini adalah perakitan tes.

B. Prosedur Penelitian

Langkah-langkah pengembangan instrumen berupa tes dalam penelitian ini adalah (1) perancangan tes, (2) uji coba tes, dan (3) perakitan tes untuk pengukuran. Pada penelitian pengembangan instrumen tes ini hanya akan dilakukan sampai pada tahap uji coba tes. Lebih lengkap akan dijelaskan di bawah ini.

1. Perancangan tes

a. Penentuan tujuan tes

Tujuan dari tes ini adalah untuk mengukur keterampilan berpikir tingkat tinggi siswa pada mata pelajaran fisika SMA Kelas X MIPA.

b. Penentuan kompetensi yang akan diujikan

Setelah menentukan tujuan tes, selanjutnya adalah menentukan kompetensi yang akan diujikan. Kompetensi yang digunakan adalah Kompetensi Inti (KI) dan Kompetensi Dasar (KD) pada mata pelajaran Fisika SMA kelas X MIPA menggunakan kurikulum 2013 Revisi 2016. Kompetensi dasar yang dipilih pada aspek pengetahuan yakni KD 3.9 pada mata pelajaran Fisika kelas X MIPA Kurikulum 2013 Revisi 2016. Dari kompetensi dasar dan Kompetensi inti yang telah ditentukan maka dibuatlah indikator yang menggunakan Kata Kerja Operasional pada level kognitif yang tinggi, yakni level C4 (menganalisis), C5 (mengevaluasi), dan C6 (Mencipta).

c. Penentuan materi yang akan diujikan

Sesuai KD yang telah ditentukan, maka materi yang diujikan pada tes ini adalah materi usaha dan energi. Materi yang digunakan mencakup, 1) Energi kinetik dan energi potensial (gravitasi dan pegas); 2) Konsep usaha (kerja); 3) Hubungan usaha (kerja) dan energi kinetik; 4) Hubungan usaha (kerja) dengan energi potensial; dan 5) Hukum kekekalan energi mekanik.

d. Penyusunan kisi-kisi tes

Untuk dapat membuat item soal yang baik diperlukan kisi-kisi tes. Pada langkah ini penyebaran soal diperhatikan berdasarkan KI, KD, materi, dan keterampilan berpikir tingkat tinggi (menganalisis, mengevaluasi, dan menciptakan) yang ingin dicapai dalam indikator yang telah disusun sebelumnya.

e. Penulisan item berdasarkan prinsip-prinsip pengembangan tes keterampilan berpikir tingkat tinggi

Item-item soal dibuat berdasarkan kisi-kisi yang telah disusun dengan memperhatikan prinsip keterampilan berpikir tingkat tinggi, yaitu dengan menampilkan stimulus berupa cerita, gambar, grafik, tabel, dll agar siswa bisa terangsang untuk berpikir sebelum mengerjakan soal yang diberikan. Adapun format butir soal yang digunakan adalah *two-tier multiple choice* atau pilihan ganda beralasan atau pilihan ganda yang membutuhkan *langkah problem solving*.

f. Penyusunan pedoman penskoran

Soal perlu dilengkapi dengan pedoman penskoran agar mudah saat memberikan skor kepada hasil tes siswa. Pedoman penskoran harus jelas dan adil, tidak ada unsur subyektif. Pedoman penskoran yang digunakan dalam penelitian ini merupakan pedoman penskoran yang dimodifikasi dari pedoman penskoran oleh Wardani dkk. (2015) sebagaimana Tabel 3.

Tabel 3. Pedoman Penskoran

	Tipe Jawaban Siswa		Skor
	<i>First-tier</i>	<i>Second-tier</i>	
Jawaban Benar	Langkah <i>Problem Solving</i> Benar dan Tuntas		3
		Langkah <i>Problem Solving</i> Benar dan Tidak Tuntas	2
Jawaban Salah	Langkah <i>Problem Solving</i> Benar dan Tuntas		2
		Langkah <i>Problem Solving</i> Benar dan Tidak Tuntas	1
Jawaban Benar	Langkah <i>Problem Solving</i> Salah		1
Jawaban Salah	Langkah <i>Problem Solving</i> Salah		0

g. Validasi item tes

Item-item soal yang telah dibuat kemudian diuji validitasnya (ketepatannya). Validasi yang dilakukan adalah validasi isi yang meliputi aspek materi, desain, dan bahasa. Validasi dilakukan oleh dosen ahli dan praktisi pendidikan yang berkompeten dalam bidang materi usaha dan energi dan pembelajarn fisika.

Data yang diperoleh dari validator berupa data kuantitatif. Data tersebut menggunakan skor skala likert yaitu 1,2 3, 4, dan 5.

Selanjutnya dianalisis total skor empirik validator dibagi dengan skor maksimal yang diharapkan melalui perhitungan (Sudjana, 2012: 100).

Tabel 4. Kriteria Penilaian untuk Hasil Validasi Ahli dan Uji Coba Produk

Indikator	Skor Pilihan Jawaban				
	5	4	3	2	1
Kelayakan Komponen materi	Sangat Layak	Layak	Cukup Layak	Kurang Layak	Tidak Layak
Kelayakan Komponen Evaluasi Pembelajaran	Sangat Layak	Layak	Cukup Layak	Kurang Layak	Tidak Layak
Kelayakan Komponen Kebahasaan	Sangat Layak	Layak	Cukup Layak	Kurang Layak	Tidak Layak

Perolehan hasil validasi instrumen tes selanjutnya dikategorikan sesuai dengan kriteria hasil evaluasi pada Tabel 5.

Tabel 5. Kriteria Hasil Evaluasi Validitas Instrumen

Nilai rata-rata	Kriteria evaluasi
25.00%– 40.00%	Tidak valid (tidak boleh digunakan)
41.00 % – 55.00%	Kurang valid (tidak boleh digunakan)
56.00% – 70.00%	Cukup valid (boleh digunakan setelah direvisi besar)
71.00% – 85.00%	Valid (boleh digunakan dengan revisi kecil)
86.00 %– 100.00%	Sangat valid (sangat baik untuk digunakan)

h. Perbaikan item dan perakitan tes

Setelah item-item tes divalidasi dan mendapat saran masukan dari validator selanjutnya item-item tersebut diperbaiki sesuai saran yang

diberikan oleh para validator. Setelah item-item soal tes diperbaiki, item-item tersebut dirakit menjadi sebuah tes.

2. Uji Coba Tes

a. Penetapan subyek uji coba

Subyek yang dipilih dalam uji coba pengembangan instrumen berupa tes keterampilan berpikir tingkat tinggi ini adalah siswa kelas X MIA MAN Pringsewu. uji coba hanya diberlakukan pada 30 % siswa kelas X MIA MAN Pringsewu.

b. Pelaksanaan uji coba tes

Uji coba dilakukan pada siswa kelas X MIA pada semester genap tahun ajaran 2018/2019.

c. Analisis data hasil uji coba

Data yang diperoleh dianalisis secara kuantitatif. Analisis data yang digunakan untuk instrumen tes berpikir tingkat tinggi ini adalah :

1) Validitas Butir soal

Validitas adalah suatu ukuran yang menunjukkan tingkat-tingkat kecermatan atau kesahihan suatu alat ukur atau instrumen dalam melakukan fungsi ukurnya. Instrumen yang valid mempunyai validitas tinggi, artinya instrumen tersebut dapat mengukur apa yang hendak diukur.

Uji validitas terhadap soal berpikir tingkat tinggi menggunakan analisis data yang dilakukan menggunakan *Model Rasch* dan dibantu oleh *Software Ministep Winsteps*. *Model Rasch* mampu

melihat interaksi antara responden dan item sekaligus. Linacre dalam Sumintono & widhiarso (2014) menjelaskan bahwa penggunaan *Model Rasch* untuk data politomi yang dikembangkan oleh Andrich tetap berlandaskan pada dua teorema dasar, yakni tingkat kemampuan atau kesetujuan individu dan tingkat kesulitan item untuk disetujui. Riswanda (2018) dalam penelitian pengembangannya menggunakan analisis *Model Rasch* untuk menentukan validitas butir soal yang dilihat dari kesesuaian soal (*item fit*) berdasarkan parameter nilai outfit dengan ketentuan sebagai berikut:

a. Nilai *Outfit Mean Square* (MNSQ) yang diterima yaitu:

$$0,5 < \text{MNSQ} < 1,5$$

b. Nilai *Outfit Z-Standard* (ZSTD) yang diterima yaitu:

$$-2,0 < \text{ZSTD} < +2,0$$

c. Nilai *Pt Mean Corr* yang diterima yaitu:

$$0,32 < \text{Pt Measure Corr} < 0,80$$

Jika butir soal memperoleh nilai yang masuk rentang nilai pada ketiga kriteria tersebut maka dikatakan valid. Sedangkan jika pada ketiga kriteria tersebut tidak terpenuhi, dapat dipastikan bahwa butir soalnya kurang bagus sehingga perlu diperbaiki atau diganti.

2) Reliabilitas

Reliabilitas mengacu pada konsistensi pengukuran, yaitu instrumen yang reliabel adalah instrumen yang tetap konsisten dan stabil dari

waktu ke waktu, dimana instrumen tersebut memiliki kehandalan sebagai alat ukur. Reliabilitas juga menunjukkan pengertian bahwa suatu instrumen cukup dapat dipercaya untuk digunakan sebagai alat pengumpul data karena instrumen tersebut sudah tepat dan sesuai.

Reliabilitas instrumen diperlukan untuk mendapatkan data yang sesuai dengan tujuan pengukuran. Oleh karena itu, dilakukan analisis data dengan menggunakan pendekatan model *Rasch* melalui program *Ministeps Winsteps*. Hasil analisis reliabilitas dapat dilihat pada program *Ministep Winsteps* dengan menggunakan *output tabel summary statistic*. Realibilitas yang ditampilkan adalah realibilitas soal (*item reliability*) dan realibilitas respon (*person reliability*). Realibilitas instrumen secara keseluruhan dapat ditentukan melalui nilai *item reliability* (Widyaningsih dan Yusuf, 2018). Adapun kriteria reliabilitas dapat dilihat pada Tabel 6.

Tabel 6. Nilai *Person* dan *Item Reliability*

Nilai <i>Reliability</i>	Keterangan
<0,6	Rendah
0,6-0,7	Cukup
0,7-0,8	Tinggi
0,8-0,9	Sangat Tinggi
>0,9	Istimewa

(Alias dkk., 2019)

Selain diperoleh realibilitas instrumen dan responden juga diperoleh nilai interaksi antara realibilitas responden dengan butir soal secara keseluruhan yang ditampilkan dalam nilai Crombach. Adapun kriteria realibilitas interaksi sesuai ketentuan Tabel 7.

Tabel 7. Nilai Crombach

Nilai <i>Cronbach's</i>	Keterangan
<0,5	Buruk
0,5 - 0,6	Kurang Baik
0,6 – 0,7	Cukup
0,7 – 0,8	Baik
>0,81	Baik Sekali

(Widyaningsih dan Yusuf, 2018)

V. SIMPULAN DAN SARAN

A. Simpulan

Instrumen tes keterampilan berpikir tingkat tinggi yang dikembangkan telah memenuhi standar kelayakan instrumen yaitu valid dan reliabel sehingga dapat digunakan untuk mengukur keterampilan berpikir tingkat tinggi siswa pada materi usaha dan energi. Hal ini dibuktikan dengan hasil analisis data uji coba yang diperoleh menggunakan model *rasch*, yaitu:

1. Semua soal keterampilan berpikir tingkat tinggi materi usaha dan energi yang dikembangkan dalam bentuk *two-tier multiple choice* telah memenuhi kriteria valid.
2. Instrumen tes yang dikembangkan memiliki realibilitas tinggi dengan nilai realibilitas sebesar 0,77.

B. Saran

Adapun saran dari penelitian pengembangan ini, yaitu:

1. Bagi guru diharapkan dapat menggunakan instrumen tes ini untuk mengukur keterampilan berpikir tingkat tinggi siswa pada materi usaha dan energi. Guru juga diharapkan mengembangkan instrumen tes sejenis berdasarkan indikator berpikir tingkat tinggi pada setiap materi

fisika lain sehingga guru dapat melatih keterampilan berpikir tingkat tinggi siswa.

2. Bagi sekolah, hasil penelitian pengembangan ini diharapkan dapat menjadi salahsatu dasar untuk mengetahui tingkat keterampilan berpikir tingkat tinggi siswa dan selanjutnya pihak sekolah dapat memfasilitasi guru untuk mengembangkan instrumen keterampilan berpikir tingkat tinggi pada materi dan pelajaran lainnya.
3. Bagi peneliti lain, diharapkan peneliti berikutnya dapat melakukan penelitian mengenai pengembangan instrumen tes berpikir tingkat tinggi pada materi fisika lainnya.

DAFTAR PUSTAKA

- Alias, M., Karim, A.A., dan Iksan Z. 2019. Kesahan dan Kebolehpercayaan Instrumen Kemahiran Menyelesaikan Masalah untuk Murid Sekolah Menengah Berdasarkan Model Pengukuran Rasch. *Jurnal Pendidikan Malaysia*, Vol. 44 No.1: hlm. 3-19. Tersedia di <http://ejournal.ukm.my/jpend/article/view/44.01SI-01>. Diakses pada Juli 2019.
- 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. Tersedia di <https://www.uky.edu/~rsand1/china2018/texts/Anderson-Krathwohl%20%20A%20taxonomy%20for%20learning%20teaching%20and%20assessing.pdf>. Diakses pada 10 April 2018.
- Arifin, Z. 2017. Kriteria Instrumen dalam Suatu Penelitian. *Jurnal THEOREMS (The Original Research of Mathematics)*, Vol.2 No. 1: hlm. 28-36. Tersedia di <http://jurnal.unma.ac.id/index.php/th/article/view/571>. Diakses pada Maret 2018.
- Cullinane, A. & Maeve, L. 2011. *Two-Tier Multiple Choiche Question: An Alternative Method Of Formatif Assesment For Fist Year Undergraduate Biology Students*. Limerick: National Center For Excellence In Mathematics and Education Science Teaching And Learning (NCEMSTL). Tersedia di https://www.academia.edu/11481815/An_Alternative_Method_of_Formative_Assessment_for_First_Year_Undergraduate_Biology_Students. Diakses pada April 2018.
- Dahlan. A. 2016. Pengemebangan Instrumen Tes Keterampilan Teknologi pada Mata Pelajaran Fisika Peserta Didik SMA di DIY. *Unnes Physic Education Journal*, Vol. 5 No. 1: hlm. 26-34. Tersedia di <https://journal.unnes.ac.id/sju/index.php/upej/article/view/12702>. Diakses pada Juni 2018.
- Depdiknas. 2007. *Peraturan Menteri Pendidikan Nasional No 20 Tahun 2007 tentang Standar Penilaian*. Jakarta: Depdiknas. hlm. 3. Tersedia di <https://luk.staff.ugm.ac.id/atur/bsnp/Permendiknas20-2007StandarPenilaian.pdf>. Diakses pada 17 Mei 2018.

- Depdikbud. 2013. *Salinan Lampiran Permendikbud No. 64 Tahun 2013 tentang Standar Isi Pendidikan Dasar dan Menengah Kurikulum 2013*. Jakarta: Depdikbud. hlm. 69. Tersedia di <http://bsnpindonesia.org/2013/06/20/permendikbud-tentang-kurikulum-tahun-2013/>. Diakses pada 10 Juli 2019.
- Dewati, D., Hardiati, D., & Fadhilah, R. 2016. Pengembangan Instrumen Penilaian Tes Diagnostik Pilihan Ganda Dua Tingkat untuk Mengukur Hasil Belajar Siswa Materi Hidrokarbon di SMA 10 Negeri Pontianak. *Ar-Razi Jurnal Ilmiah*, Vol. 4 No. 2: hlm. 23-32. Tersedia di <http://openjurnal.unmuhpkn.ac.id/index.php/ar-r/article/view/671>. Diakses pada Juli 2019.
- Djaali. 2004. *Pengukuran dalam Bidang Pendidikan*. Jakarta: Program Pascasarjana Universitas Negeri Jakarta. hlm. 10-11.
- Ekawati, F., Handhika, J., dan Huriawati, F. 2017. Etnosains dan Peranannya dalam Menguatkan Karakter Bangsa. *Prosiding Seminar Nasional Pendidikan Fisika III 2017*. ISSN : 2527-6670. Tersedia di <http://e-journal.unipma.ac.id/index.php/snpf/issue/view/148>. Diakses pada September 2019.
- Elviyanti, M.Y. 2018. Pengembangan Instrumen Penilaian Berpikir Tingkat Tinggi Pada Mata Pelajaran Bahasa Indonesia, Semester I, Kelas XI. *Proceeding of The 1st International Conference on Education Language And Literature (ICON-ELITE)*. Tersedia di <http://icon-elite.conference.unesa.ac.id>. Diakses pada 20 Juni 2019.
- Fanani, M.Z. 2018. Strategi Pengembangan Soal *Higher Order Thinking Skill* (HOTS) dalam Kurikulum 2013. *Journal of Islamic Religious Education*, Vol. 2 No. 1: hlm. 57-76. Tersedia di <https://jurnal.iainkediri.ac.id/index.php/edudeena/article/view/582>. Diakses pada Agustus 2019.
- Greenstein, L. 2012. *Assessing 21st Century Skills*. United State of America: Corwin a Sage Company. Tersedia di <http://us.corwin.com/en-us/nam/assessing-21st-century-skills/book237748>. Diakses pada 12 Juni 2018.
- Heong, Y.M., Othman, W.D., Md Yunos, J., Kiong, T.T., Hassan, R., & Mohamad, M.M. 2011. The Level of Marzano Higher Order Thinking Skills Among Technical Education Students. *International Journal of Social and Humanity*, Vol. 1 No. 2: pp. 121-125. Tersedia di https://www.researchgate.net/publication/255720579_The_Level_of_Marzano_Higher_Order_Thinking_Skills_among_Technical_Education_Students. Diakses pada 15 Juli 2018.

- Himah, F., Sudarti, & Subiki. 2016. Pengembangan Instrumen Tes Computer Based Test-Higherorder Thinking (CBT-HOT) Pada Mata Pelajaran Fisika Di SMA. *Jurnal Pembelajaran Fisika*, Vol. 5 No. 1: hlm. 89-95. Tersedia di <https://jurnal.unej.ac.id/index.php/JPF/article/view/3569>. Diakses pada 10 Juni 2018.
- Iskandar, D. dan Senam. 2015. Studi Kemampuan Guru Kimia SMA Lulusan UNY dalam Mengembangkan Soal UAS Berbasis HOTS. *Jurnal Inovasi Pendidikan IPA*, Vol. 1 No.1: hlm. 65-72. Tersedia di <https://journal.uny.ac.id/index.php/jipi/article/view/4533>. Diakses pada 20 Oktober 2018.
- Istiyono, E., Djemari, M., dan Suparno. 2014. Pengembangan Tes Keterampilan Berpikir Tingkat Tinggi Fisika (Pyshots) Peserta Didik SMA. *Jurnal Penelitian dan Evaluasi Pendidikan*, Vol. 1: hlm. 1-12. Tersedia di <https://journal.uny.ac.id/index.php/jpep/article/view/2120>. Diakses pada Januari 2018.
- Ivanty, D.W.N., Aminah, N.S., dan Ekawati, E.Y. 2013. Penyusunan Instrumen Tes Tengah Semester Genap Fisika X SMA untuk Kelas X SMA. *Jurnal Pendidikan Fisika*, Vol. 1 No.1: hlm. 27-36. Tersedia di <http://jurnal.fkip.uns.ac.id/index.php/pfisika/article/view/1789>. Diakses pada Agustus 2019.
- Johnson, E.B. 2007. *Contextual Teaching & Learning, Menjadikan Kegiatan Belajar Mengajar Mengasyikan Dan Bermakna* (diterjemahkan oleh Ibnu Setiawan). Bandung: Penerbit MLC. hlm. 69. Tersedia di <http://opac.perpusnas.go.id/DetailOpac.aspx?id=1119550>. Diakses pada 20 Mei 2018.
- Kadir, A. 2015. Menyusun dan Menganalisis Tes Hasil Belajar. *Jurnal Al-Ta'dib*, Vol. 8 No.2: hlm. 70-81. Tersedia di <http://ejournal.iainkendari.ac.id/al-tadib/article/view/411>. Diakses pada 20 Juni 2018.
- Kemendikbud. 2017. Modul Penyusunan Soal Higher Order Thinking Skills (HOTS). Jakarta: Dirjen Pendidikan Dasar dan Menengah.
- Lewy, Zulkardi, & Aisyah, N. 2009. Pengembangan Soal Untuk Mengukur Keterampilan Berpikir Tingkat Tinggi Pokok Bahasan Barisan Dan Deret Bilangan Di Kelas IX Akselerasi SMP Xaverius Maria Palembang. *Jurnal Pendidikan Matematika*, Vol. 3 No. 2: hlm. 14-28. Tersedia di <https://ejournal.unsri.ac.id/index.php/jpm/article/view/326>. Diakses pada 20 Juni 2018.
- Malik, A., Ertikanto, C., & Suyatna, A. 2015. Deskripsi Kebutuhan HOTS Assessment Pada Pembelajaran Fisika Dengan Metode Inkuiri Terbimbing. *Prosiding Seminar Nasional Fisika (E-Journal) SNF2015*, e-ISSN: 2476-9398. Tersedia di

<http://journal.unj.ac.id/unj/index.php/prosidingsnf/article/download/6618/4752/>. Diakses pada 10 Juni 2018.

- Matondang, Zulkifli. 2009. Validitas dan realibilitas suatu instrumen penelitian. *Jurnal Taburasa PPS Unimed*. Vol. 6 No. 1: hlm. 87-79. Tersedia di <http://digilib.unimed.ac.id/705/1/Validitas%20dan%20reliabilitas%20suatu%20instrumen%20penelitian.pdf>. Diakses pada Januari 2018.
- Napitupulu, D. 2017. Analisis Model Rasch Pada Instrumen Keberterimaan Marka Optik Penanda Jarak Aman Kendaraan. *Jurnal Penelitian Transportasi Darat*, Vol. 19 No.1: hlm.37-48. Tersedia di <http://ojs.balitbanghub.dephub.go.id/index.php/jurnalدارat/article/view/604>. Diakses pada Februari 2018.
- Noviana, M., Sajidan, dan Karyanto, P. 2014. Pengembangan Instrumen Evaluasi *Two-Tier Multiple Choice Question* untuk Mengukur Keterampilan Berpikir Tingkat Tinggi. *Jurnal Pendidikan Sains*, Vol. 1 No. 2: hlm. 1-12. Tersedia di <http://jurnal.fkip.uns.ac.id/index.php/snps/article/view/5088>. Diakse pada 25 September 2019.
- Novinda, M.R.R. 2019. Pengembangan Tes Pilihan Ganda Menggunakan Model *Rasch* Materi Gerak Lurus Kelas X Pontianak. *Jurnal Untan*, Vol. 8 No. 6: hlm. 1-11. Tersedia di <http://jurnal.untan.ac.id/index.php/jpdpb/article/view/33452>. Diakses pada 10 Oktober 2019.
- Purba, S.E.D. 2018. Analisis model *Rasch* instrumen tes prestasi pada mata pelajaran dasar dan pengukuran listrik. *Jurnal Penelitian dan Evaluasi Pendidikan*, Vol. 6 No. 2: hlm. 142-147. Tersedia di <http://jurnal.ustjogja.ac.id/index.php/wd/article/view/3393>. Diakses pada Oktober 2019.
- Ramadhan, G., Dwijanati, P., dan Wahyuni, S. 2018. Analisis Kemampuan Berpikir Tingkat Tinggi (*High Order Thinking Skills*) Menggunakan Instrumen *Two Tier Multiple Choice* Materi Konsep dan Fenomena Kuantum Siswa SMA di Kabupaten Cilacap. *Unnes Physics Education Journal*, Vol. 7 No. 3: hlm. 85-90. Tersedia di <https://journal.unnes.ac.id/sju/index.php/upej/article/view/27682>. Diakses pada Oktober 2019.
- Riswanda, J. 2018. Pengembangan Soal Berbasis *Higher Order Thinking Skill* (HOTS) serta Implementasinya di SMA Negeri 8 Palembang. *Didaktika Biologi*, Vol. 2 No. 1: hlm. 49–58. Tersedia di <http://jurnal.umpalembang.ac.id/dikbio/article/view/1648/1385>. Diakses pada Oktober 2019.
- Rofiah, E., Aminah, N.S., & Ekawati, E.Y. 2013. Penyusunan Instrumen Tes Keterampilan Berpikir Tingkat Tinggi Fisika pada Siswa SMP. *Jurnal*

- Pendidikan Fisika*, Vol. 1 No. 2: hlm. 17-22. Tersedia di <http://www.jurnal.fkip.uns.ac.id/index.php/pfisika/article/view/2797>. Diakses pada Maret 2018.
- Saputro, B., & Supahar. 2018. Pengembangan Instrumen Penilaian Kemampuan Berpikir Tingkat Tinggi untuk Mengukur Pencapaian Hasil Belajar Fisika Peserta Didik SMA Kelas XI Materi Optika. *Jurnal Pendidikan Fisika*, Vol. 7 No. 6: hlm.1-6. Tersedia di <http://journal.student.uny.ac.id/ojs/index.php/pfisika/article/view/11431>. Diakses pada Oktober 2019.
- Sari, A.L.R., Parno dan Ahmad Taufiq. 2016. Kemampuan Berpikir Kritis dan Pemahaman Konsep Fisika Siswa SMA pada Materi Hukum Newton. *Prosiding Seminar Nasional Pendidikan IPA Pascasarjana UM*. Vol 1. ISBN 978-602-9286-21-2. Tersedia di <http://pasca.um.ac.id/wp-content/uploads/2017/02/Ayu-Lingga-88-99.pdf>. Diakses pada Mei 2018.
- Sayuti, A.S. 2010. *Metode Penelitian Pendidikan*. Jakarta: Grafindo. hlm. 19.
- Shidiq, A.S., Masykuri, M., dan Susanti, V.H. 2014. Pengembangan Penilaian Two-Tier Multiple Choiche Untuk Mengukur Kemampuan Berpikir Tingkat Tinggi(Higher Order Thinking Skills) Pada Materi Kelarutan Untuk Siswa SMA/MA Kelas XI. *Jurnal Pendidikan Kimia*. Vol. 3 No. 4: hlm. 83-92. Tersedia di <http://jurnal.fkip.uns.ac.id/index.php/kimia/article/view/4507>. Diakses pada September 2019.
- Siswoyo dan Sunaryo. 2017. High Order Thinking Skills: Analisis Soal dan Implementasinya dalam Pembelajaran Fisika di Sekolah Menengah Atas. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, Vol. 3 No.1: hlm. 11-20. Tersedia di <http://journal.unj.ac.id/unj/index.php/jpppf/article/view/2498>. Diakses pada Juli 2018.
- Sudaryono, Margono, G., & Rahayu, W. 2013. *Pengembangan Instrumen Penelitian Pendidikan*. Yogyakarta: Graha Ilmu. hlm. 45.
- Sudjana, N. 2012. *Penilaian Hasil Blajar Mengajar*. Bandung: PT Remaja Rosdakarya. hlm. 100.
- Sugiarti, T., I Kaniawati, & L Aviyanti. 2017. Development of Assesment Instrumen of Critical Thingking in Physics at Senior High School. *Journal of Physics*, Vol. 812 No. 12018: pp. 1-8. Tersedia di <https://ui.adsabs.harvard.edu/abs/2017JPhCS.812a2018S/abstract>. Diakses pada Mei 2018.
- Sugiyono. 2010. *Metode Penelitian Kuantitatif Kualitatif & RND*. Bandung: Alfabeta. hlm. 32.

- Sumintono, B., dan Widhiarso, W. 2015 *Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial*. Cimahi: Trikomunikata. hlm. 75.
- Suryani, Y.E. 2018. Aplikasi *Rasch Model* dalam Mengevaluasi *Intelligenz Structure Test (IST)*. *Jurnal Penelitian Psikologi*, Vol. 3 No. 1: hlm. 73–100. Tersedia di <http://journal.walisongo.ac.id/index.php/Psikohumaniora/article/view/2052>. Diakses pada September 2019.
- Trilling, B., & Fadel, C. 2009. 21st Century Skills. *Learning for life in our times*. Jossey-Bass. San Fransisco, California. Tersedia di <https://psynet.apa.org/record/2009-18745-000>. Diakses pada 19 Mei 2018.
- Wardani, R.K., Yamtinah, S., dan Mulyani, B. 2015. Instrumen Tes Penilaian Two Tier Test Aspek Pengetahuan Untuk Mengukur Tingkat Keterampilan Proses Sains(KPS) Pada Pembelajaran Kimia Untuk Siswa SMA/MA Kelas X. *Jurnal Pendidikan Kimia*. Vol. 4 No. 4: hlm. 156-162. Tersedia di <http://jurnal.fkip.uns.ac.id/index.php/kimia/article/view/6639/4530>. Diakses pada Mei 2018.
- Widyaningsih, S.W. dan Yusuf, I. 2018. Analisis Soal Modul Laboratorium Fisika Sekolah I Menggunakan *Racsh Model*. *Jurnal Gravity: Jurnal Ilmiah Penelitian dan Pembelajaran Fisika*, Vol. 4 No. 1: hlm. 33-46. Tersedia di <http://jurnal.untirta.ac.id/index.php/Gravity>. Diakses pada 10 April 2018.
- Yusmanto, H., Soetjipto, B.E., & Djatmika, E.T. 2016. *Higher Order Thinking Skills* Siswa SMPS IT Darul Azhar Berdasarkan Taksonomi Bloom Revisi. *Jurnal Pendidikan Humaniora*, Vol. 6 No. 1: hlm. 1-13. Tersedia di <http://pasca.um.ac.id/conferences/index.php/gtk/article/view/198>. Diakses pada September 2019.