

## II. KAJIAN PUSTAKA

### A. Kajian Teori

#### 1. Penilaian dalam Pembelajaran Fisika

Assesmen atau penilaian diartikan sebagai kegiatan menafsirkan data hasil pengukuran berdasarkan kriteria maupun aturan-aturan tertentu (Widoyoko, 2012: 3). Penilaian juga diartikan sebagai suatu kegiatan untuk mengumpulkan informasi tentang proses dan hasil belajar peserta didik dalam rangka membuat keputusan-keputusan berdasarkan kriteria dan pertimbangan tertentu yang dilakukan secara sistematis dan berkesinambungan (Arifin, 2016: 4). Adanya proses penilaian menjadi hal yang penting dalam pembelajaran karena dapat memberikan banyak manfaat baik bagi peserta didik maupun pendidik untuk memperbaiki kualitas dari pembelajaran tersebut.

Kesesuaian butir soal dengan indikator penilaian sangat penting untuk membangun kesimpulan validitas tes standar, karena pada akhirnya, indikator kemudian dapat digunakan untuk menyesuaikan pengajaran dengan akuntabilitas (Anderson, Irvin, Alonzo, et al., 2015: 22). Menurut Faize & Dahar (2012: 3), tes adalah media untuk menilai pemahaman peserta didik mengenai pembelajaran. Tes standar merupakan tes yang prosedur, administrasi, bahan dan aturan penskorannya tetap sehingga sejauh mungkin penilaian sama pada tempat dan waktu yang berbeda (Nitko & Brookhart, 2011: 347-348). Hal senada dinyatakan oleh Kara & Celikler (2015: 21) bahwa tes adalah alat penilaian yang digunakan untuk menentukan

kemampuan peserta didik yang berkaitan dengan domain kognitif dalam pendidikan

Sudjana (2016: 3) mengemukakan beberapa fungsi dari penilaian dalam pembelajaran, yaitu (a) Alat untuk mengetahui tercapai-tidaknya tujuan instruksional. Dengan fungsi ini maka penilaian harus mengacu kepada rumusan-rumusan tujuan instruksional. (b) Umpan balik bagi perbaikan proses belajar-mengajar. Perbaikan mungkin dilakukan dalam hal tujuan instruksional, kegiatan belajar peserta didik, strategi mengajar guru, dan lain-lain. (c) Dasar dalam menyusun laporan kemajuan belajar peserta didik kepada para orang tuanya. Dalam laporan tersebut dikemukakan kemampuan dan kecakapan belajar peserta didik dalam berbagai bidang studi dalam bentuk nilai-nilai prestasi yang dicapainya.

Kara & Celikler (2015: 21) menjelaskan tujuan utama dari penilaian dan evaluasi adalah (a) Untuk menentukan tingkat kemampuan siswa kejenjang berikutnya. (b) Untuk menentukan berapa banyak siswa memiliki perilaku yang akan diajarkan dalam kursus. (c) Untuk menentukan berapa banyak siswa mencapai keuntungan dari program dan seberapa banyak pembelajaran terjadi. (d) Memberi tahu siswa tentang kekurangan mereka. (e) Untuk menilai keterampilan siswa. (f) Untuk memotivasi siswa untuk mengikuti kursus dan bidang belajar yang berbeda. (g) Mengevaluasi keabsahan program pendidikan atau metode program.

Penilaian tidak hanya digunakan sebagai alat untuk melihat ketercapaian kemampuan peserta didik. Penilaian juga memiliki peran yang penting untuk

mengarahkan peserta didik tentang apa yang harus dilakukan setelah pembelajaran. Dengan demikian tujuan penilaian hendaknya diarahkan pada empat hal berikut (Kusaeri & Suprananto, 2012: 9) (a) Penelusuran (*keeping track*), yaitu untuk menelusuri agar proses pembelajaran tetap sesuai dengan rencana. (b) Pengecekan (*checking-up*), yaitu untuk mengecek adakah kelemahan-kelemahan yang dialami oleh siswa selama proses pembelajaran. (c) Pencarian (*findingout*), yaitu mencari dan menemukan hal-hal yang menyebabkan terjadinya kelemahan dan kesalahan dalam proses pembelajaran. (d) Penyimpulan (*summing-up*), yaitu untuk menyimpulkan apakah siswa telah menguasai seluruh kompetensi yang ditetapkan dalam kurikulum atau belum.

Dari tujuan tersebut, guru sebagai pendidik juga dapat mengambil manfaat dari penilaian yang dilakukan pada peserta didik. Agar penilaian juga bermakna untuk guru, maka seorang guru harus menguasai beberapa pengetahuan terkait dengan penilaian pendidikan dalam kegiatan pembelajaran, serta harus memiliki beberapa kemampuan (Kusaeri & Suprananto, 2012: 17), diantaranya (a) Mampu memilih prosedur-prosedur penilaian yang tepat untuk membuat keputusan pembelajaran. (b) Mampu mengembangkan prosedur penilaian yang tepat untuk membuat keputusan pembelajaran. (c) Mampu dalam melaksanakan, melakukan penskoran, serta menafsirkan hasil penilaian yang telah dibuat. (d) Mampu menggunakan hasil-hasil penilaian untuk membuat keputusan-keputusan di bidang pendidikan. (e) Mampu mengembangkan prosedur penilaian yang valid dan

menggunakan informasi penilaian. (f) Mampu dalam mengkomunikasikan hasil penilaian.

Penilaian merupakan bagian penting dari pendidikan. Tidak hanya dapat digunakan sebagai alat untuk mengetahui pencapaian hasil belajar peserta didik, tetapi juga dapat digunakan sebagai dasar bagi pendidik untuk membuat langkah pembelajaran selanjutnya. Penilaian yang baik tidak hanya dapat mengumpulkan data yang memperlihatkan kemajuan belajar peserta didik, tetapi juga dapat digunakan sebagai bahan evaluasi pembelajaran baik bagi peserta didik maupun bagi pendidik agar didapat kualitas pembelajaran yang lebih baik.

Sukiman (2012: 18) menambahkan beberapa ciri dari evaluasi pembelajaran, yaitu (a) Evaluasi yang dilakukan untuk mengukur keberhasilan belajar dilakukan secara tidak langsung. (b) Pengukuran penilaian hasil belajar dilakukan secara kuantitatif. (c) Evaluasi hasil belajar digunakan sebagai unit yang tetap. (d) Prestasi belajar yang dicapai bersifat relatif. (e) Evaluasi hasil belajar tidak akan terlepas dari kekeliruan pengukuran.

Ciri tersebut dapat digunakan oleh pendidik sebagai acuan dalam mengembangkan penilaian yang baik. Agar penilaian menjadi lebih bermakna baik untuk peserta didik maupun pendidik, maka pemilihan penilaian yang akan digunakan haruslah sesuai. Kesesuaian ini dapat bergantung pada tujuan penilaian, jenis penilaian, maupun dari aspek atau pertimbangan lainnya.

## 2. Keterampilan Berpikir Tingkat Tinggi Versi Marzano

Keterampilan berpikir tingkat tinggi didefinisikan sebagai penggunaan pikiran secara lebih luas untuk menemukan tantangan baru. Keterampilan berpikir tingkat tinggi ini menghendaki seseorang untuk menerapkan informasi baru atau pengetahuan sebelumnya dan memanipulasi informasi untuk menjangkau kemungkinan jawaban dalam situasi yang baru (Heong, et al, 2011: 122). Berpikir tingkat tinggi tidak hanya diartikan sebagai sekedar menghafal suatu fakta atau pengetahuan saja, tetapi lebih kepada mengeksplorasi pengalaman baik untuk menganalisis, mensintesis, ataupun mengevaluatif. King, et al, (2012: 34) menyatakan bahwa seseorang melakukan analisis, sintesis, dan evaluasi dengan tujuan untuk memahami fakta, konsep-konsep, prinsip-prinsip, dan prosedur atau langkah-langkah yang ada. Jadi, keterampilan berpikir tingkat tinggi (*higher order thinking skills*) memberikan peserta didik tentang keterampilan hidup yang relevan dan menawarkan mereka sebuah manfaat tambahan untuk membantu mereka meningkatkan pengetahuan yang dimiliki.

Yen & Halili (2014: 41) berupaya untuk memperjelas dimensi *higher order thinking* yang terlihat membantu untuk referensi di masa mendatang untuk guru dan peneliti serupa. Ia merangkum pengetahuan untuk mengajar berpikir menjadi “pengetahuan tentang unsur-unsur pemikiran” bersama-sama dengan empat sub-kategori. Pertama, pengetahuan tentang strategi berpikir individu (membuat perbandingan, merumuskan argumen yang benar, menarik kesimpulan yang valid, dan lain-lain). Kedua, pengetahuan tentang

genre pemikiran (argumentasi, belajar penyelidikan, pemecahan masalah, berpikir kritis, berpikir ilmiah, berpikir kreatif, dan lain-lain). Ketiga, pengetahuan tentang metakognisi (berpikir tentang pemikiran sendiri). Keempat, pengetahuan tentang isu-isu tambahan (disposisi berpikir (habits of mind), budaya berpikir, dan lain-lain).

Tujuan utama pendidikan sains adalah untuk membantu peserta didik mengembangkan kemampuan berpikir tingkat tinggi agar mereka dapat menghadapi tantangan dalam kehidupan sehari-hari (Saido, Siraj, Nordin, & Amedy, 2015: 13). Pengetahuan dan kompetensi sangat penting untuk menjamin kualitas pelaksanaan higher order thinking skills (Sulaiman, Muniyan, Madhvan, Hasan, & Rahim, 2017: 1). Melalui pertanyaan HOTS mampu melatih peserta didik untuk berpikir kreatif, kritis dan inovatif (Ping, Ahmad, Adnan, & Hua, 2017:1). Karakter inilah yang nantinya akan menjadi bekal untuk peserta didik dalam kehidupan yang sebenarnya.

HOTS dianggap sebagai tingkat yang signifikan karena individu dapat berkontribusi terhadap kemajuan pendidikan yang menghasilkan solusi untuk masalah yang ada dan menerapkan konsep baru ke struktur yang ada (Nagayar, Ahmad, & Kanniah, 2015: 53). Hal tersebut dapat dijumpai dari berfikir tingkat rendah terlebih dahulu. Low order thinking skills penting dalam menyediakan platform dasar untuk penerapan higher order thinking skills (Kamarulzaman, Sailin, Mahmor, & Shaari, 2017: 71). Kemampuan berpikir bukanlah hal instan yang diperoleh tanpa melalui proses.

Soal HOTS tidak selalu berupa soal yang sulit untuk dikerjakan. Soal HOTS memiliki beberapa karakteristik yang membedakan dengan soal LOTS. Karakteristik HOTS yaitu: (1) Mengukur kemampuan berfikir tingkat tinggi. (2) Berbasis permasalahan kontekstual (berupa kasus), (3) Tidak rutin atau tidak akrab, (4) Menggunakan bentuk soal yang beragam dan memiliki stimulus yang menarik (Fanani, 2018: 57).

Guru harus berperan untuk menumbuhkan penggunaan pertanyaan yang menuntut penggunaan pemikiran tingkat tinggi di kelas sambil mengenali kesulitan yang dihadapi peserta didik dalam memecahkan masalah yang berkaitan dengan HOTS (Abdullah, Abidin, & Ali, 2015: 140). Program HOTS, berdasarkan teori pengolahan informasi, memisahkan keterampilan berpikir ke dalam kategori berdasarkan bagaimana peserta didik memproses informasi (Eisenman & Payne, 2017: 5). Informasi ini nantinya dimanfaatkan oleh peserta didik maupun pendidik untuk memperbaiki kualitas pembelajaran.

Ada banyak jenis keterampilan berpikir tingkat tinggi yang berkembang, mulai dari HOTS itu sendiri hingga dari taksonomi yang ada, seperti Piaget, Bruner, Bloom, Gagné, dan Marzano. Penelitian Marzano (Heong et al, 2011: 122) mengidentifikasi 13 keterampilan berpikir tingkat tinggi versi Marzano, yaitu abstraksi, membuat keputusan, penalaran induktif, penalaran deduktif, membandingkan, mengklasifikasikan, analisis kesalahan, mengkonstruksi dukungan, analisis perspektif, investigasi, pemecahan masalah, penyelidikan

eksperimental, penemuan. Definisi dari masing-masing keterampilan berpikir tingkat tinggi dijelaskan dalam Tabel 1.

**Tabel 1. Indikator HOTS Versi Marzano**

No	Aspek	Sub Aspek	Definisi
1	Pemulihan Pengetahuan	Abstraksi	Menekankan pada tema yang mendasari
		Membuat keputusan	Menggunakan kriteria untuk memilih (dari berbagai pilihan)
2	Pemahaman	Penalaran Induktif	Merumuskan prinsip-prinsip dari bukti yang ada
		Penalaran Deduktif	Menerapkan prinsip-prinsip untuk menggambarkan kesimpulan
3	Analisis	Membandingkan	Menekankan persamaan dan perbedaan
		Mengklasifikasi	Mengelompokkan sesuatu ke dalam kategori
		Analisis Kesalahan	Mengkritisi pemikiran (cara berpikir)
		Mengonstruksi dukungan	Mendukung pernyataan
		Analisis Perspektif	Menekankan pandangan personal tentang isu
4	Pemanfaatan pengetahuan	Investigasi	Memanggil informasi
		Pemecahan Masalah	Menanggulangi/menyelesaikan masalah
		Pengalaman	Berusaha mencari penjelasan
		Penemuan	Mendapatkan cara yang terbaik

Berdasarkan indikator yang ada pada taksonomi Marzano, seseorang dapat diukur apakah sudah menggunakan keterampilan berpikir tingkat tingginya atau belum. Berbeda dengan taksonomi Bloom yang masih mengukur kemampuan berpikir tingkat rendah pada ranah C1-C3, dan baru mengukur kemampuan berpikir tingkat tinggi pada ranah C4-C6, taksonomi Marzano sudah sepenuhnya mengukur kemampuan berpikir tingkat tinggi sehingga



semua indikatornya dapat digunakan untuk mengukur kemampuan berpikir tingkat tinggi.

Taksonomi Marzano memiliki kesetaraan dengan Taksonomi Bloom, yaitu abstraksi (mengabstraksi/C6), membuat keputusan (memutuskan/C5), penalaran induktif (menguraikan/C4), penalaran deduktif (menyimpulkan/C5), membandingkan (membandingkan/C4)), mengklasifikasikan (klasifikasi teori/C6), analisis kesalahan (menganalisis/C4), mengkonstruksi dukungan (mengorganisasikan/C4), analisis perspektif (menganalisis/C4), investigasi (mengecek/C5), pemecahan masalah (memutuskan/C5), penyelidikan eksperimental (menelaah/C6), penemuan (mencipta/C6).

### **3. *Computerized Adaptive Test (CAT)***

Salah satu penilaian berbasis komputer adalah *Computerized Adaptive Test (CAT)*. CAT adalah pengujian berbasis komputer dimana butir soal yang diberikan kepada peserta tes disesuaikan dengan kemampuan peserta tes (Ozyurt, Ozyurt, & Baki, 2013: 2926-2927). Menurut Reckase (2009: 311) CAT adalah suatu metodologi untuk sebuah tes, menyajikan kepada peserta tes, dan menghitung skor tes menggunakan teknologi komputer interaktif. CAT menjadi bentuk penilaian yang tidak memaksakan penggunaanya mengerjakan soal yang tidak sesuai dengan dirinya.

Soal dalam CAT dapat berupa soal pilihan ganda. Menurut Cetin, Dindar, & Geban (2011: 600) tes pilihan ganda lebih disukai di kelas sains karena mudah untuk menerapkan dan mengevaluasi pemahaman peserta didik tentang subjek terkait. Untuk menghasilkan butir tes pilihan ganda, ada tiga

langkah yang diperlukan. Pertama, dibuat sebuah model struktur kognitif. Kedua, butir yang dikembangkan menggunakan isi dari model struktur kognitif. Ketiga, dihasilkan butir soal menggunakan aplikasi komputer. Berdasarkan ketiga langkah tersebut, ratusan bahkan ribuan butir dapat dihasilkan dengan menggunakan model butir tunggal (Gierl, Lai, & Turner, 2012: 759).

Butir pilihan ganda unik karena, berlawanan dengan pendapat umum, butir ini memungkinkan untuk mengukur perilaku di tingkat yang lebih tinggi dari taksonomi tujuan pendidikan (Kubiszyn & Borich, 2013: 138). Hal senada dinyatakan oleh Kastner & Stangl (2011: 265) bahwa pertanyaan pilihan ganda memiliki ragam dan daftar kemungkinan jawaban dari mana pengambil tes harus memilih jawabannya, pertanyaan pilihan ganda yang memiliki satu pilihan benar dan beberapa pilihan alternatif lain yang salah yang disebut distraktor atau pengecoh. Soal pilihan ganda inilah yang nantinya akan dikembangkan

*Computerized Adaptive Test* adalah jenis tes berbasis komputer yang menyesuaikan dengan tingkat kemampuan dari para peserta (Rezaie & Golshan, 2015: 128). Larson & Madsen (2011: 32) menyatakan bahwa CAT memberikan solusi yang menjanjikan untuk menyediakan tes dengan ketelitian yang memadai dan ukuran yang umum untuk mengukur perbedaan kemampuan individu. Berbeda dengan tes klasik, penilaian individu menjadi kenyataan melalui penilaian adaptif yang memberikan kesempatan untuk mengetahui peserta didik sesuai dengan tingkat kemampuan mereka (Ozyurt,

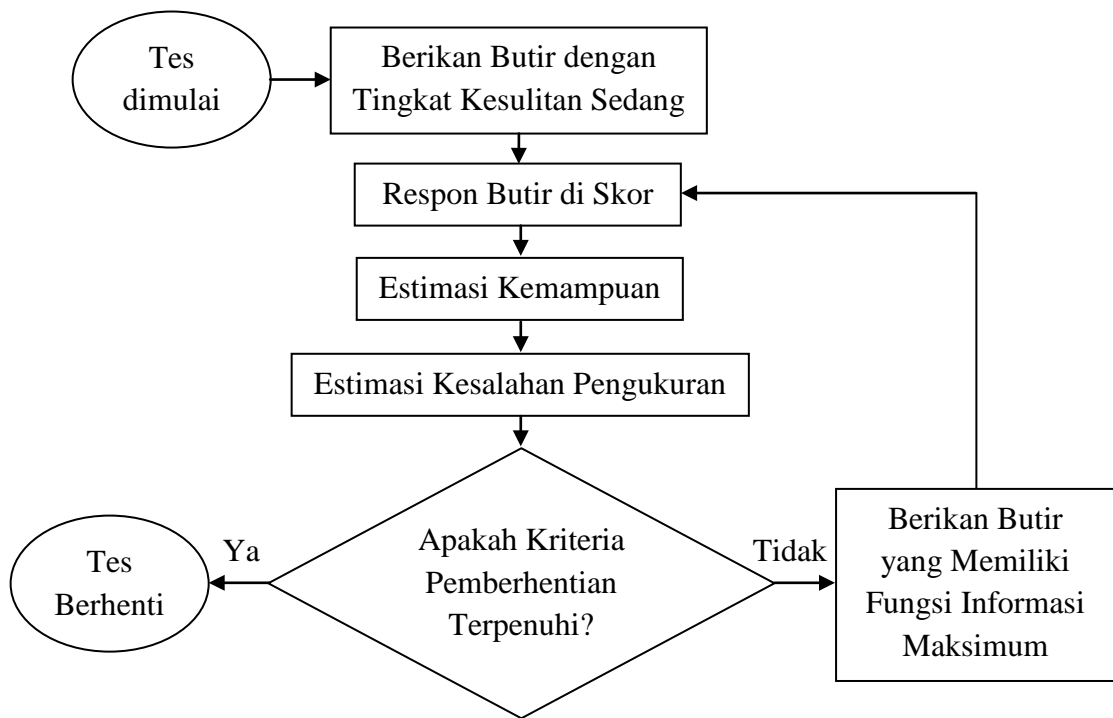
Ozyurt, Baki, & Guven, 2012: 9837). Ini menunjukkan bahwa meskipun soal yang dikerjakan berbeda, akan tetapi hal yang akan diukur tetaplah sama.

*Computerized adaptive test* merupakan sebuah metode penyampaian ujian yang canggih. Hampir selama 40 tahun telah dilakukan penelitian tentang metode penilaian ini (Thompson & Weiss, 2011: 1). CAT dapat digunakan untuk mempromosikan pembelajaran yang lebih efektif dengan menguji berbagai keterampilan, pengetahuan dan pemahaman. (Jamil, Tariq, & Shami, 2013: 371). Penilaian akan menjadi lebih bermakna karena manfaatnya bisa dirasakan dalam jangka panjang baik oleh peserta didik maupun pendidik.

Dalam penggunaan *computerized adaptive test* terdapat langkah-langkah (algoritma). Algoritma CAT meliputi beberapa langkah (Čisar, Radosav, Markoski, Pinter, & Čisar, 2010: 141) sebagai berikut:

- a. Semua butir yang belum diberikan dievaluasi untuk menentukan mana yang terbaik yang akan diberikan selanjutnya mengingat tingkat kemampuan yang diperkirakan saat ini.
- b. Butir "terbaik" berikutnya diberikan dan peserta ujian merespon.
- c. Perkiraan kemampuan baru dihitung berdasarkan tanggapan terhadap semua butir yang diberikan.
- d. Langkah a sampai c diulang sampai kriteria penghentian terpenuhi.

Berikut ditampilkan bagan alur pengujian algoritma tes adaptif pada Gambar 1.



Gambar 1. Bagan Alur Pengujian Algoritma Tes Adaptif

#### 4. Karakteristik Tes

Karakteristik penilaian didasarkan pada analisis data yang dilakukan pada beberapa aspek, yakni: (a) validasi isi (b) kecocokan item instrumen, (c) reliabilitas, (d) kurva karakteristik item (ICC), (e) indeks kesukaran, dan (f) fungsi informasi dan SEM.

##### a. Validasi Isi

Validasi soal meliputi validasi isi dilakukan oleh para ahli sesuai bidang keahlian yaitu; dosen ahli materi, dosen ahli pengukuran dan guru fisika. Validasi isi ini berdasarkan kesesuaian indikator soal terhadap soal yang ditulis. Validasi isi meliputi empat kategori yaitu tidak sesuai, kurang sesuai, sesuai dan sangat sesuai yang masing-masingnya diberi skor 1, 2, 3 dan 4. Langkah selanjutnya adalah menghitung indeks Aiken

dari masing-masing butir yang telah divalidasi. Analisis dapat digunakan untuk menentukan indeks validitas isi instrumen tes yaitu formula Aiken's seperti di bawah ini (Azwar, S., 2015:113), kemudian lembar validasi dari validator dianalisis menggunakan rumus:

$$V = \frac{s}{n(c-1)} \quad (1)$$

Keterangan:

V= Indeks Aiken;  $s = r - I_0$ ; r = angka yang diberikan penilai;  $I_0$  = angka penilaian validasi terendah; n = jumlah penilai; c = angka penilaian validasi tertinggi

Langkah kedua adalah indeks V Aiken tiap butir angket validitas dikonversikan menjadi data kualitatif dengan rentang indeks V antara 1 sampai 0. Hasil validasi dinyatakan valid apabila indeks V Aiken memiliki nilai dengan rentang 0,74 sampai 1,00 sesuai dengan jumlah *raters* dan jumlah kategori.

#### **b. Uji Kecocokan Item Instrumen (*goodness of fit*)**

Analisis data pada uji coba bertujuan untuk mengetahui validitas dan reliabilitas seluruh maupun setiap butir soal yang dikembangkan. Analisis ini bertujuan untuk memperlihatkan pengujian *goodness of fit* tes secara keseluruhan maupun per item soal. Pada proses pengujian fit tes secara keseluruhan yang telah dikembangkan menurut Adam & Khoo (1996: 30) berdasarkan nilai rerata INFIT Mean of Square (Mean INFITMNSQ) beserta simpangan bakunya atau mengamati nilai rata INFIT t (Mean INFIT t) beserta simpangan bakunya.

### c. Reliabilitas

Output program Quest tidak hanya dapat digunakan untuk menguji kecocokan, tetapi juga dapat digunakan untuk menampilkan *reliabilitas of estimates* set instrumen tes. Wright (2008: 146-147) menyatakan bahwa reliabilitas adalah sebuah statemen tentang kekonsistenan dan stabilitas dari skor sebuah instrumen. Berdasarkan hasil analisis dengan acuan nilai interpretasi instrumen menggunakan Rasch Model (Sumintono & Widhiarso, 2015: 109) dapat dikategorikan seperti pada Tabel 2:

**Tabel 2. Nilai Interpretasi Reliabilitas PCM**

Nilai Reliability	Interpretasi Reliabilitas
>0,94	Istimewa
0,91 – 0,94	Bagus sekali
0,81 – 0,90	Bagus
0,67 – 0,80	Cukup
<0,67	Lemah

Berdasarkan hasil analisis diperoleh indeks kesukaran atau tingkat kesukaran ( $b$ ) untuk masing-masing item. Item dikatakan baik jika indeks kesukaran berada diantara -2,0 sampai dengan 2,0 yang dapat dinyatakan dengan  $(-2,0 < b < 2,0)$ . Karena Indeks kesulitan butir sangat rendah, jika nilai  $b_i$  mendekati -2 dan indeks kesulitan butir sangat tinggi untuk suatu kelompok peserta tes jika nilai  $b_i$  mendekati +2. Semakin besar nilai parameter  $b_i$ , maka semakin besar kemampuan yang diperlukan untuk menjawab benar.

#### **d. Kurva Karakteristi Item (ICC)**

Seluruh item soal memiliki karakteristik masing-masing, untuk menunjukkan karakteristik item soal itu maka di gunakan bantuan program Parscale. Dengan menggunakan bantuan program Parscale akan menampilkan keluaran berbentuk Kurva Karakteristik Item (ICC) yang akan digunakan untuk menganalisis hasil dari uji coba tersebut.

#### **e. Fungsi Informasi dan *Standard Error of Measurement* (SEM)**

Berdasarkan analisis dengan Parscale diperoleh fungsi informasi dan *Standard Error of Measurement* (SEM). Berdasarkan fungsi informasi dan SEM, maka tes ini cocok untuk siswa dengan kemampuan ( $\theta$ ) rendah, sedang atau tinggi.

Penelitian dan pengembangan ini menghasilkan instrumen tes *Higher Order Thinking Skills* (HOTS) Fisika versi Marzano Sekolah Menengah Atas dalam bentuk empat paket soal yang berjumlah 40 butir soal tiap paket. Data hasil uji coba lapangan menghasilkan produk 136 butir soal yang valid dan reliabel. Instrumen tes Physics HOTS Marzanoian yang valid dan reliabel.

### **B. Materi Fisika**

Pengembangan tes yang dilakukan adalah materi fisika SMA kelas XI semester ganjil. Adapun kompetensi dasar dan materi pembelajaran diuraikan dalam tabel berikut.

**Tabel 3. Materi Pembelajaran Fisika Kelas XI Semester Ganjil**

Kompetensi Dasar	Materi Pembelajaran
3.1 Menerapkan konsep torsi, momen inersia, titik berat, dan momentum sudut pada benda tegar (statis dan dinamis) dalam kehidupan sehari-hari 4.1 Membuat karya yang menerapkan konsep titik berat dan keseimbangan benda tegar	Keseimbangan dan dinamika rotasi: <ul style="list-style-type: none"> <li>• Momen gaya</li> <li>• Momen inersia</li> <li>• Keseimbangan benda tegar</li> <li>• Titik berat</li> <li>• Hukum kekekalan momentum sudut pada gerak rotasi</li> </ul>
3.2 Menganalisis sifat elastisitas bahan dalam kehidupan sehari-hari 4.2 Melakukan percobaan tentang sifat elastisitas suatu bahan berikut presentasi hasil dan makna fisisnya	Elastisitas dan Hukum Hooke: <ul style="list-style-type: none"> <li>• Hukum Hooke</li> <li>• Susunan pegas seri-paralel</li> </ul>
3.3 Menerapkan hukum-hukum fluida statik dalam kehidupan sehari-hari. 4.3 Merencanakan dan melakukan percobaan yang memanfaatkan sifat-sifat fluida statis, berikut presentasi hasil dan makna fisisnya	Fluida statik: <ul style="list-style-type: none"> <li>• Hukum utama hidrostatis</li> <li>• Tekanan Hidrostatis</li> <li>• Hukum Pascal</li> <li>• Hukum Archimedes</li> <li>• Meniskus</li> <li>• Gejala kapilaritas</li> <li>• Viskositas dan Hukum Stokes</li> </ul>
3.4 Menerapkan prinsip fluida dinamik dalam teknologi 4.4 Membuat dan menguji proyek sederhana yang menerapkan prinsip dinamika fluida, dan makna fisisnya	Fluida Dinamik: <ul style="list-style-type: none"> <li>• Fluida ideal</li> <li>• Azas kontinuitas</li> <li>• Azas Bernoulli</li> <li>• Penerapan Azas Kontinuitas dan Bernoulli dalam Kehidupan</li> </ul>
3.5 Menganalisis pengaruh kalor dan perpindahan kalor yang meliputi karakteristik termal suatu bahan, kapasitas, dan konduktivitas kalor pada kehidupan sehari-hari 4.5 Merencanakan dan melakukan percobaan tentang karakteristik termal suatu bahan, terutama terkait dengan kapasitas dan konduktivitas kalor, beserta presentasi hasil dan makna fisisnya	Suhu, Kalor dan Perpindahan Kalor: <ul style="list-style-type: none"> <li>• Suhu dan pemuaian</li> <li>• Hubungan kalor dengan suhu benda dan wujudnya</li> <li>• Azas Black</li> <li>• Perpindahan kalor secara konduksi, konveksi, dan radiasi</li> </ul>

### C. Kajian Penelitian yang Relevan

Dalam penelitian Kamaruddin & Haryanto (2014: 39-40) menyatakan bahwa penggunaan perangkat pengujian tes berbasis *Computerized Adaptive Test* (CAT) mampu menyajikan item tes secara acak sehingga dapat mengurangi kecurangan



pada waktu ujian berlangsung. Materi yang dikembangkan dalam tes tersebut hanya satu materi, yaitu menganalisis rangkaian listrik, sedangkan materi yang digunakan dalam penelitian dan pengembangan ini meliputi lima materi, yaitu kesetimbangan dan dinamika rotasi, elastisitas, fluida statis, fluida dinamis, serta suhu dan kalor.

Istiyono, et al (2014: 1) melakukan penelitian untuk mempelajari kegunaan dan keefektifan soal pilihan ganda beralasan untuk mengukur keterampilan berpikir tingkat tinggi mata pelajaran fisika SMA pada implementasi kurikulum 2013. Hasil penelitian ini menunjukkan bahwa soal pilihan ganda beralasan pada mata pelajaran fisika SMA valid dan reliabel sehingga dapat digunakan untuk mengukur keterampilan berpikir tingkat tinggi secara efektif dan efisien. Soal ini dapat digunakan pada peserta didik dalam jumlah yang banyak dengan biaya yang terjangkau dan waktu yang dibutuhkan untuk mengoreksi relatif cepat. Penelitian tersebut relevan dengan penelitian yang akan dilakukan. Hal yang membedakan adalah tujuan dari penelitian tersebut. Penelitian yang sudah dilakukan bertujuan untuk mencari seberapa efektif penggunaan CAT, sedangkan penelitian yang akan dilakukan bertujuan untuk mencari efektifitas dan pemetaan penggunaannya.

Rezaie & Golshan (2015: 133) menyatakan bahwa CAT yang berbasis sistem adaptif dengan pemberian soal sesuai kemampuan peserta didik dapat membuat proses tes lebih efektif, efisien, fleksibel, dan individual, individual dalam arti soal yang diberikan berbeda beda sesuai kemampuan peserta didik. CAT tersebut digunakan pada pengujian tes bahasa, sedangkan CAT yang dikembangkan merupakan pengujian pada tes fisika. Penelitian tersebut sebatas mencari

keuntungan dan kerugian dari CAT, sedangkan pengembangan CAT yang dilakukan memiliki tujuan yang lebih spesifik yaitu untuk memetakan kemampuan berpikir.

Penelitian dari Ermansah & Istiyono (2017: 80) menunjukkan bahwa instrumen tes Physics Higher Order Thinking Skills (HOTS) Marzanoian memiliki karakteristik kelayakan digunakan untuk mengukur. Hal ini didasarkan pada indeks V Aiken yang dikategorikan baik. Selain itu, hasil uji coba menyatakan fit dengan model Partial Credit Model (PCM) dengan data Polytomus empat kategori. Penelitian ini dilakukan pada kelas X, sedangkan pengembangan yang dilakukan pada penelitian ini ditujukan untuk kelas XI, begitu juga dengan materi soal yang dikembangkan

Penelitian dan pengembangan dari Kusuma, Rosidin, Abdurrahman, & Suyatna (2017: 26) dilakukan untuk mengembangkan butir soal berupa soal HOTS. Soal HOTS tersebut dibuat untuk meningkatkan kemampuan berpikir ilmiah. Soal yang dibuat didasarkan pada indikator Taksonomi Marzano, yaitu C4 (menganalisis), C5 (mengevaluasi), dan C6 (membuat). Penelitian yang akan dilakukan dibedakan berdasarkan taksonomi yang digunakan, yaitu Taksonomi Marzano yang terdiri dari 4 aspek dan 13 subaspek indikator.

Penelitian yang dilakukan Wu, Kuo & Wang (2017: 61) menerapkan CAT dalam penilaian pembelajaran matematika. Hasil dari penelitian ini menunjukkan bahwa CAT dapat meningkatkan keefektifitasan proses penilaian. Penelitian tersebut mengembangkan CAT dengan menggunakan tes berupa pilihan ganda biasa yang terdiri dari empat pilihan jawaban, sedangkan pada penelitian CAT-

MARZANO jenis tes yang digunakan adalah tes dengan pilihan ganda beralasan (Reasoning Multiple Choice) dengan lima pilihan jawaban dan lima pilihan alasan untuk mengukur dan memetakan HOTS Fisika versi Marzano peserta didik.

#### **D. Kerangka Berpikir**

Penilaian dalam pembelajaran merupakan hal yang sangat penting dilakukan pendidik untuk melihat keberhasilan peserta didik dalam memahami pembelajaran. Penilaian memungkinkan pendidik untuk mengevaluasi efektivitas pembelajaran dan kemudian membuat perbaikan pada metode penyampaian, kegiatan pembelajaran atau sumber belajar. Penilaian merupakan hal yang penting, karena secara tidak langsung penilaian juga mempengaruhi proses sebuah pembelajaran. Oleh karena itu, cara dan jenis penilaian harus sesuai dengan aspek yang akan dinilai atau diukur. Jika penilaiannya tepat dan sesuai, maka akan ada banyak sekali informasi yang diperoleh dari penilaian tersebut. Tidak terkecuali penilaian dalam pembelajaran fisika. Terlebih fisika merupakan ilmu fundamental yang menjadi payung besar perkembangan berbagai ilmu pengetahuan dan teknologi.

Pembelajaran fisika merupakan sekumpulan pengetahuan yang unik. Pembelajaran fisika meliputi proses berpikir, cara berpikir, sebuah bentuk penyelidikan dan interaksi antara teknologi dan sosial dalam memahami fisika. Untuk mengukur kemampuan fisika perlu dilakukan proses penilaian yang mampu mengukur kemampuan berpikir tingkat rendah dan tinggi peserta didik. Karena hal yang akan diukur selalu berbeda, maka alat penilaian yang digunakan

untuk menilai juga pasti berbeda. Perbedaan alat penilaian ini selalu disesuaikan dengan karakteristik yang akan dinilai.

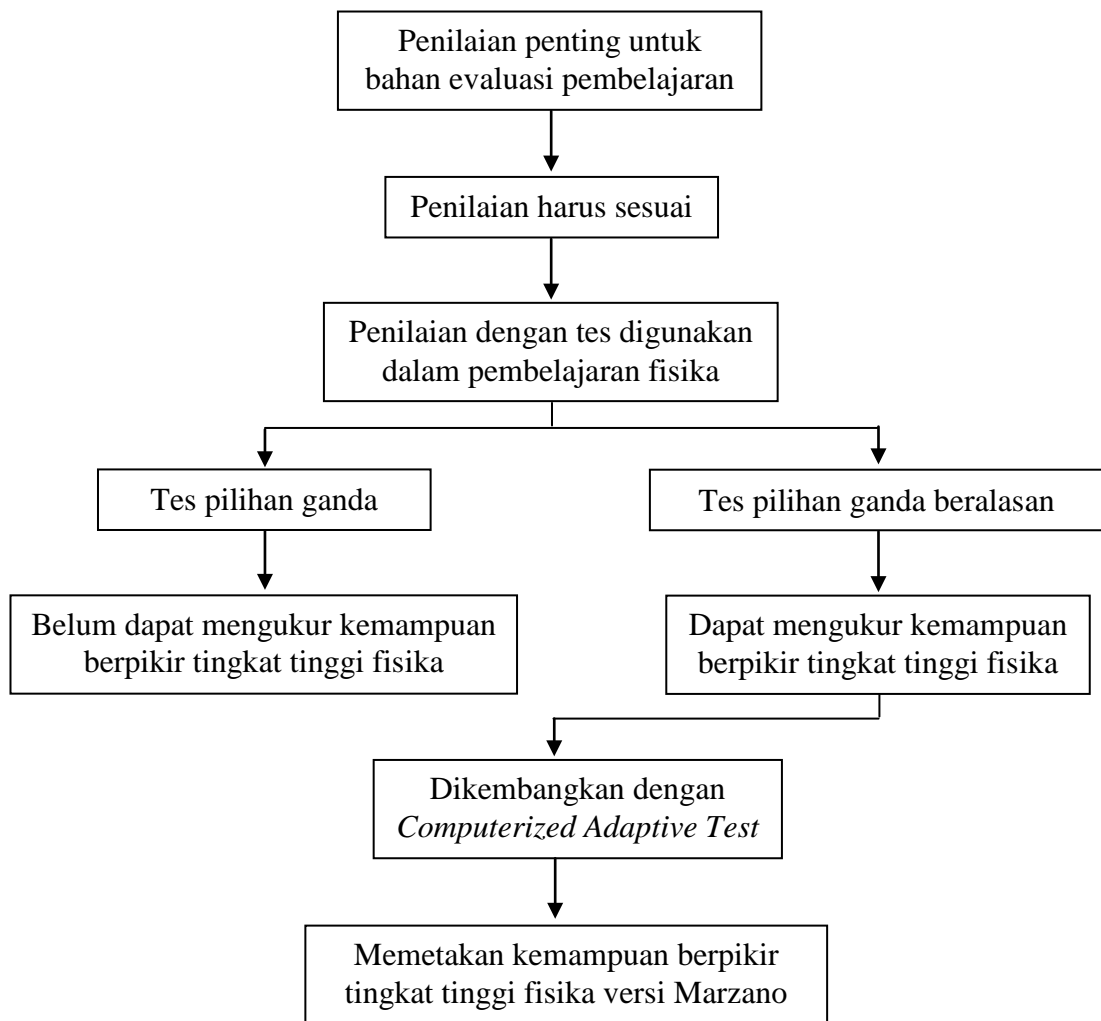
Alat penilaian yang digunakan dalam pembelajaran fisika meliputi tes dan non tes. Selain sebagai alat penilaian digunakan untuk mengukur hasil belajar peserta didik, penilaian juga memiliki peran lain dalam pembelajaran. Penilaian juga berperan sebagai proses untuk memperoleh informasi yang digunakan agar dapat mengambil keputusan tentang peserta didik, kurikulum sekolah dan kebijakan pendidikan. Hal tersebut menjelaskan bahwa penilaian merupakan hal sangat penting bagi dunia pendidikan sebagai salah satu cara memperbaiki mutu atau kualitas dari pendidikan itu sendiri.

Salah satu jenis penilaian yang ada adalah tes. Tes dapat dikelompokkan ke dalam dua bentuk yaitu tes objektif dan tes non objektif. Tes objektif meliputi pilihan ganda (majemuk), jawaban singkat, menjodohkan, dan uraian objektif, sedangkan bentuk soal non objektif adalah uraian bebas, unjuk kerja atau observasi, dan portofolio atau proyek. Masing-masing jenis tes tersebut dapat berkembang sesuai dengan kebutuhan dari penggunaan tes tersebut. Tes pilihan majemuk beralasan (*Reasoning Multiple Choice*) misalnya, sebagai contoh bentuk pengembangan dari tes pilihan majemuk biasa. Tes tersebut dikembangkan untuk mendapatkan informasi lebih dari peserta didik, yaitu alasan mengapa peserta didik memilih jawaban tersebut. Tes yang akan dikembangkan pada penelitian ini berupa tes yang dapat mengukur kemampuan berpikir tingkat tinggi (HOTS) versi Marzano.

Tes yang banyak digunakan di kalangan sekolah hanya berupa tes pilihan ganda yang belum dapat mengukur semua kemampuan berpikir tingkat tinggi, selain itu penggunaan kertas dan pensil masih mendominasi dalam tes tersebut. Seiring dengan berkembangnya ilmu pengetahuan dan teknologi, komputer sudah banyak digunakan dalam kegiatan belajar mengajar, akan tetapi masih belum banyak komputer yang digunakan sebagai media tesnya. Terlebih tes berbasis komputer yang dapat menyesuaikan dengan kemampuan berpikir peserta didik. Melalui *Computerized Adaptive Test* (CAT) diharapkan peserta didik dapat mengerjakan soal-soal sesuai dengan kemampuannya, terutama kemampuannya dalam berpikir tingkat tinggi.

Sesuai dengan HOTS versi Marzano, keterampilan berpikir tingkat tinggi meliputi 13 indikator, yaitu: abstraksi, membuat keputusan, penalaran induktif, penalaran deduktif, membandingkan, mengklasifikasikan, analisis kesalahan, mengkonstruksi dukungan, analisis perspektif, investigasi, pemecahan masalah, penyelidikan eksperimental, penemuan. Berdasarkan 13 indikator inilah soal-soal akan dibuat dalam bentuk tes yang dapat beradaptasi dengan berbantuan komputer. Peserta didik diharapkan dapat memiliki keterampilan berpikir tingkat tinggi dengan menggunakan tes adaptasi berbasis komputer.

Setelah tes *higher order thinking skills* fisika versi Marzano berbasis *computerized adaptive test* dibuat, langkah selanjutnya adalah memetakan kemampuan berpikir peserta didik. Kemampuan berpikir tingkat tinggi merupakan *output* utama yang dijadikan acuan keberhasilan pembelajaran peserta didik di sekolah. Bagan kerangka berpikir secara sederhana disajikan dalam Gambar 2.



Gambar 2. Bagan Kerangka Berpikir

### E. Pertanyaan Penelitian

1. Bagaimana konstruksi tes HOTS Fisika versi Marzano yang dikembangkan berdasarkan materi fisika SMA Kelas XI semester gasal?
2. Bagaimana konstruksi tes HOTS Fisika versi Marzano yang dikembangkan berdasarkan aspek dan subaspek pada ranah kognitif HOTS Marzano?
3. Bagaimana validasi isi tes HOTS Fisika versi Marzano yang dikembangkan?

4. Bagaimana kecocokan item instrumen tes HOTS Fisika versi Marzano yang dikembangkan?
5. Bagaimana reliabilitas tes HOTS Fisika versi Marzano yang dikembangkan?
6. Bagaimana kurva karakteristik item (ICC) tes HOTS Fisika versi Marzano yang dikembangkan?
7. Bagaimana indeks kesukaran tes HOTS Fisika versi Marzano yang dikembangkan?
8. Bagaimana fungsi informasi dan *standard error measurement* tes HOTS Fisika versi Marzano yang dikembangkan?
9. Bagaimana profil kemampuan HOTS Fisika versi Marzano menggunakan *computerized adaptive test* pada peserta didik kelas XI di kabupaten Bantul?
10. Bagaimana tingkat keberhasilan *computerized adaptive test* yang dikembangkan untuk memetaan tes HOTS Fisika versi Marzano?