

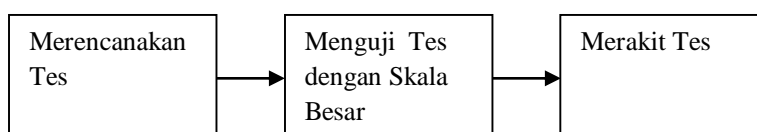
BAB III

METODE PENELITIAN

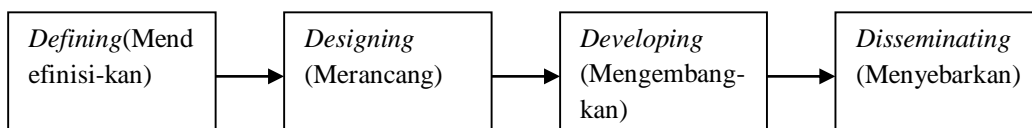
A. Model Pengembangan

Pengembangan instrumen tes menggunakan modifikasi Model Wilson dan Model Oriondo dan Antonio (1998), dengan langkah yang dilakukan adalah sebagai berikut: (1) perancangan tes, (2) uji coba tes, dan (3) perakitan tes. (Istiyono, Mardapi, & Suparno, 2013).

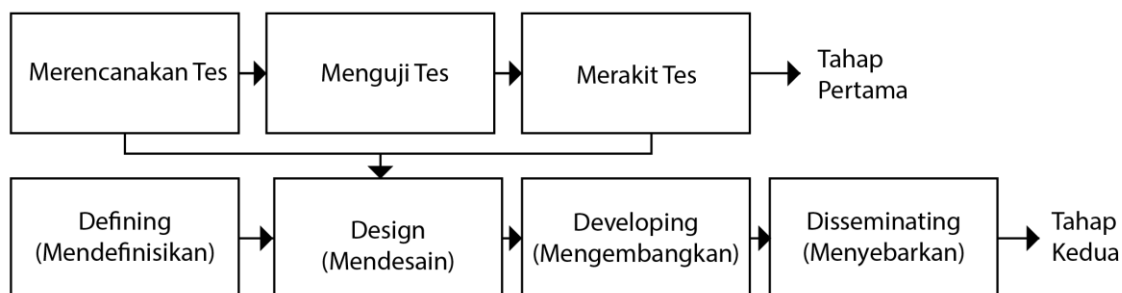
Model pengembangan yang digunakan untuk mengembangkan CAT adalah 4-D Thiagarajan, Samuel dan Semmel yang meliputi empat tahap yaitu *Defining*, *Designing*, *Developing* dan *Disseminating* (Hobri, 2010) .



Gambar 2a. Model Wilson Oriondo dan Antonio



Gambar 2b. Model 4D dari Thiagarajan, Samuel dan Semmel



Gambar 3. Modifikasi Model Pengembangan

B. Prosedur Pengembangan

Tahapan pengembangan dalam penelitian dilakukan sesuai model dan dibagi kedalam 2 tahap, tahap pertama adalah tahap pengembangan instrumen tes seperti Gambar 2, kemudian tahap kedua adalah pengembangan media CAT sesuai Gambar 3 Hal ini dilakukan karena pengembangan media CAT memerlukan masukan berupa instrumen tes, sehingga pengembangan instrumen tes dilakukan terlebih dahulu.

1. Merencanakan Tes

Pada tahapan pertama pengembangan instrumen tes dilakukan perencanaan terhadap instrumen tes yang meliputi:

- a. Menentukan tujuan tes yang disesuaikan dengan tujuan pembelajaran
- b. Menentukan kompetensi inti dan kompetensi dasar yang sudah menjadi kebijakan kurikulum.
- c. Menentukan materi yang diujikan, yakni materi fisika kelas XI SMA semester gasal menyesuaikan kurikulum 2013 edisi revisi.
- d. Menyusun matriks soal sebagai panduan dan informasi jumlah butir soal yang akan dikembangkan.
- e. Menyusun kisi-kisi soal yang merupakan acuan soal itu dibuat. Kisi-kisi soal disusun dengan memperhatikan aspek dan sub aspek dalam kemampuan berpikir kreatif (elaborasi, fleksibilitas, kelancaran, dan originalitas).

- f. Menulis butir yang terdiri dari 5 pilihan jawaban dan 5 pilihan alasan. Hal ini disebut juga sebagai tes pilihan ganda beralasan (*two-tier multiple choice*).
- g. Menyusun pedoman penskoran diperlukan agar tidak terjadi penskoran yang berbeda antara satu dengan yang lain. Sistem penskoran menggunakan model Rasch 1PL yakni yaitu; *pertama* kriteria nilai (1) jika jawaban salah dan alasan salah, *Kedua* kriteria nilai (2) jika jawaban benar dan alasan salah, *ketiga* kriteria nilai (3) jika jawaban salah dan alasan benar dan *keempat* kriteria nilai (4) jika jawaban benar dan alasan benar.
- h. Menguji validitas butir oleh ahli, yakni validasi yang dilakukan oleh ahli dari pendidikan fisika.
- i. Melakukan perbaikan butir dari hasil revisi validator
- j. Melakukan perakitan tes.

2. Menguji Tes dengan Skala Besar

Setelah tes terkonstruksi, perlu dilakukan uji skala besar untuk mendapatkan data, adapun langkahnya sebagai berikut:

- a. Menetapkan subjek uji coba di beberapa SMA di Gunung Kidul dengan jumlah sampel minimal sebanyak 250 peserta didik (Sion, 2009; 27). Pemilihan sampel menggunakan metode pengambilan sampel acak berstrata.
- b. Melaksanakan uji coba dilakukan kurang lebih pada bulan 3 tahun 2018.

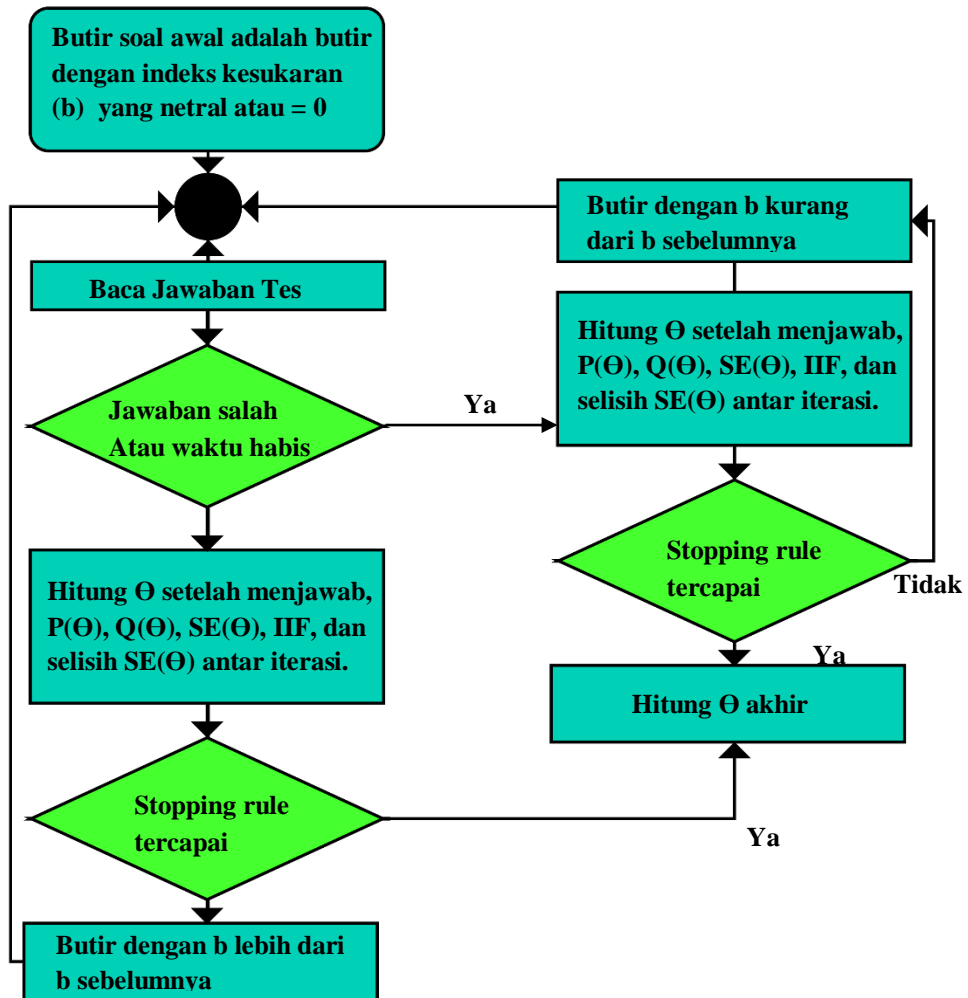
- c. Melakukan analisis dari hasil dari uji terbatas dengan bantuan program *Quest*, *Parscale* dan *excel*, untuk memperoleh informasi karakteristik butir.

3. Perbaikan dan Perakitan Tes

Penelitian dilanjutkan dengan melakukan perakitan tes yang telah diuji sebelumnya coba kedalam media CAT. Butir yang sudah divalidasi dan diuji kemudian dipilah dan diperbaiki menjadi sebuah instrumen tes.

4. Pendefinisian CAT (*Define*)

Tahap ini meliputi analisis studi literatur terkait pengembangan tes penggunaan CAT sebagai media pemberian tes. Pemilihan media dengan *computer adaptive test* (CAT) karena menurut kajian literatur memiliki banyak keunggulan, diantaranya cocok untuk meningkatkan efektifitas dan efisiensi tes sumatif, dapat diaplikasikan dengan teori respon butir, dan memberikan motivasi kepada peserta didik untuk memaksimalkan kemampuan mereka karena tes yang mereka kerjakan akan beradaptasi terhadap jawaban sebelumnya. Selanjutnya dianalisis alur penyusunan algoritmanya dan dibuat diagram alir sebagai panduan.



Gambar 4. Diagram Alir CAT

5. Perancangan CAT (Design)

Desain atau perancangan meliputi pembuatan desain visual dan algoritma media. Langkah yang ditempuh adalah sebagai berikut:

a. Pengembangan Acuan dan Kriteria Media

Pengujian terhadap kelayakan media menurut acuan dan kriteria dengan cara merancang garis besar dalam pembuatan *blue print*.

b. Pemilihan Format Media

Pemilihan media adalah dengan komputer. Hal ini dilakukan karena fleksibilitas komputer sebagai media untuk menampung data guru, peserta, tes, dan lainnya.

c. Desain awal Media

Pengembangan media dengan web berbasis server dan klien. Hal penting yang harus diperhatikan adalah pembuatan *database* dan desain visualisasi CAT.

6. Pengembangan CAT (Development)

Pada tahapan ini dihasilkan instrumen yang sudah divalidasi dan direvisi berdasarkan masukan dari para ahli. Soal yang disiapkan secara garis besar melalui dua kegiatan, yaitu *expert appraisal* dan *development testing*.

a. Validasi Ahli Media

Validasi ini disebut juga *expert appraisal*, prosesnya dilakukan oleh ahli media. *Developmental testing* merupakan kegiatan uji coba rancangan produk. Pada uji coba ini dicari respon, reaksi atau komentar dari pengguna CAT. Informasi ini akan digunakan untuk memperbaiki produk untuk kemudian diujikan dalam skala besar untuk memperoleh hasil yang efektif.

b. Revisi Produk

Revisi dilakukan berdasarkan saran para ahli dan hasil uji coba media yang siap digunakan untuk tahapan uji lapangan.

7. Penyebaran CAT (*Disseminate*)

Produk yang dikembangkan disebarkan terlebih dahulu ke sekolah yang diujikan. Hal ini sebagai apresiasi terhadap pihak sekolah yang sudah mengizinkan. Selanjutnya produk akan dikembangkan dengan cara digabungkan dengan pengukuran berpikir tingkat tinggi yang lain. Hasil penelitian akan disebarkan dengan cara diterbitkan kedalam jurnal penelitian dan diseminarkan melalui seminar internasional.

C. Desain Uji Coba Produk

1. Desain Uji Coba

Desain uji coba produk meliputi validasi dan pengemasan soal menjadi 4 paket soal yang siap diuji coba. Tahap pelaksanaan uji terbatas dengan mengambil sampel di beberapa SMA di Gunung Kidul. Hasil uji terbatas berupa validitas soal dan tingkat kesulitan butir soal untuk dimasukkan ke dalam CAT. Setelah diperoleh informasi tingkat kesukaran, butir disusun ke dalam CAT untuk dilakukan uji lapangan ke beberapa sekolah.

2. Subjek Uji Coba

Subjek uji terbatas dalam penelitian ini adalah peserta didik kelas XI dari SMA di Kabupaten Gunung Kidul, Yogyakarta. Sementara uji lapangan ditargetkan 3 kelas,. Yakni di SMAN 1 Wonosari, SMAN 1 Karangmojo dan SMAN 1 Tanjungsari. Peserta tes uji lapangan tidak sebesar uji terbatas, karena tidak ada syarat batas peserta.

3. Teknik dan Instrumen Pengumpulan Data

a. Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam pengembangan tes *berpikir kreatif* fisika untuk kelas XI meliputi tes dan non tes. Teknik tes dilakukan untuk mengukur berpikir kreatif peserta didik. Dengan teknik tes ini dipetakan kemampuan berpikir kreatif peserta didik di kabupaten Gunung Kidul. Teknik non tes yang dilakukan untuk mengetahui validasi instrumen dan media serta keterlaksanaan CAT meliputi efektifitas penggunaannya.

b. Instrumen Pengumpulan Data

Penelitian ini menggunakan empat paket instrument tes kemampuan berpikir kreatif, angket validasi media (validasi CAT), angket validasi instrumen tes berpikir kreatif fisika dan angket keterlaksanaan CAT. Instrumen tes masing-masing paket terdiri dari 50 soal dengan 10 anchor, uji validasi instrumen dan media, kemudian angket efektifitas yang meliputi kebermanfaatan, kepraktisan, dan kelayakan CAT.

4. Teknik Analisis Data

Teknik analisis data dilakukan untuk mengolah data menjadi informasi, sehingga karakteristik data tersebut menjadi mudah untuk dipahami. Dalam pengembangan ini analisis data dilakukan untuk mengetahui karakteristik dan konstruksi instrumen yang dikembangkan.

a. Validitas Isi

Validitas isi dilakukan untuk untuk mengetahui keterukuran yang dilakukan oleh butir instrumen. Data penelitian dianalisis secara deskriptif. Pemberian lembar validasi dosen ahli untuk mengetahui kualitas produk dan instrumen yang dikembangkan. Angket validitas disusun dengan skala 1 sampai 4. Rincian skala yang dilakukan adalah 1 “sangat tidak sesuai”, 2 “tidak sesuai”, 3 “Sesuai”, 4 “Sangat Sesuai”. Selanjutnya data kuantitatif ini dirubah menjadi kualitatif dan dianalisis dengan indeks V Aiken. Analisis dilakukan oleh Azwar (2012, 139) sebagai berikut :

$$V = \sum s / [n(c-1)] \quad (1)$$

Keterangan: V= Indeks Aiken

I_0 = Skala Terkecil

r = dari I_0+1 sampai $I_0+(c-1)$

s = r- I_0

Langkah selanjutnya mengonversikan menjadi data kualitatif untuk masing-masing butir dengan rentang indeks V antara 1 sampai 0. Arikunto (2006; 276) menyatakan hasil validasi dinyatakan valid apabila indeks V Aiken memiliki nilai dengan rentang 0,76 sampai 1,00.

b. Kelayakan Media

Melakukan uji validitas media untuk mengetahui kelayakan media yang dikembangkan, yakni Computerized Adaptive Test. Analisis data yang dilakukan adalah mengkonversi lembar validasi ahli media menjadi

beberapa interval. Hasilnya menurut Akbar (2013) dapat dikategorikan ke dalam Tabel 5.

Tabel 5. Analisis Data Kelayakan Media

Persentase Nilai Rata-Rata	Kategori	Keterangan
85% - 100%	Sangat Layak	Sangat Baik
69% - 84%	Layak	Boleh digunakan dengan revisi kecil
53% - 68%	Cukup Layak	Boleh digunakan dengan revisi besar
37% - 52%	Kurang Layak	Tidak boleh digunakan
20% - 36%	Tidak Layak	Tidak boleh digunakan

Perhitungan persentase uji analisis pengembangan sebagai berikut:

$$P = \frac{\Sigma x}{\Sigma y} \times 100\%$$

Keterangan: P = persentase kelayakan

Σx = jumlah keseluruhan jawaban responden

Σy = jumlah skor maksimal

Berdasarkan kriteria kelayakan media oleh Akbar (2013: 78)

c. Validitas Empirik

Instrumen tes yang dinyatakan valid, diuji cobakan secara empirik kepada peserta didik. Butir tes dianalisis menggunakan teori modern item Response theory model Rasch (1PL). Penskoran butir tes menggunakan teknik *Partial Credit Model* (PCM) karena dapat mengakomodasi penilaian untuk butir politomi. PCM ini merupakan pengembangan 1-PL dari Model Rasch. Hasil uji coba dianalisis menggunakan program Quest untuk mengetahui karakteristik.

Analisis dengan program QUEST dilakukan dengan menuliskan hasil uji skala besar pada program microsoft excel. Hasil uji skala besar ini kemudian di ekspor ke dalam format text atau .Dat sesuai format yang dibutuhkan program Quest. Dengan menggunakan koding yang ada pada lampiran 2c, data tersebut diolah dan diperoleh beberapa informasi untuk dilakukan analisis sebagai berikut:

1. Uji Kecocokan Butir (*Goodness of fit*)

Analisis ini bertujuan untuk memperlihatkan pengujian *goodness of fit* tes secara keseluruhan maupun per butir. Dari olah informasi dengan program QUEST diambil informasi uji kecocokan butir berdasarkan nilai rerata INFIT *Mean of Square* (Mean INFITMNSQ) beserta nilai simpangan bakunya (Mean INFIT).

2. Tingkat Kesukaran

Indeks kesukaran (b) berlaku untuk tiap butir. Sebuah butir dikarakan baik apabila indeks kesukarannya berada pada rentang $-2,0 < b < 2,0$. Apabila kurang atau melebihi nilai tersebut, maka butir akan dibuang. Pengambilan nilai kesukaran tiap butir dari analisis QUEST diperoleh dari informasi keluaran ekstensi *idt.out*.

3. Fungsi Informasi dan SEM

Pengambilan informasi dilakukan dengan bantuan program Parscale melalui analisis untuk memperoleh fungsi informasi dan *Standard Error of Measurement* (SEM). Analisis dilakukan dengan memasukkan hasil penelitian yang sudah dikonversi kedalam file *.txt*

atau DAT. Dengan data tersebut, Parscale dapat memberikan hasil analisis yang digunakan untuk mengetahui tingkat tes ini sesuai untuk kemampuan (Θ) tertentu.

4. *Item Characteristic Curve (ICC)*

Karakteristik dan tingkat butir dapat dilihat dalam bentuk kurva. Analisis untuk mendapatkan ICC diperoleh dengan menggunakan bantuan Parscale. Pengambilan informasi ini sama langkahnya dengan pengambilan informasi fungsi informasi dan SEM.

5. Reliabilitas Instrumen

Pengukuran reliabilitas dilakukan dengan bantuan program Quest yang mengacu kepada internal consistency. Pada program Quest keluaran sh menunjukkan reliabilitas setiap orang. Sedangkan nilai reliabilitas keseluruhan dapat dilihat pada data keluaran dengan akhiran tn, pada internal konsistensi. Tes dikategorikan memiliki reliabilitas baik menurut tabel Guilford (1956) dalam Sukiman (2012) sebagai berikut:

Tabel 6. Kategori Reliabilitas Guilford

Koefisien Reliabilitas	Kategori
0,00 - 0,19	Sangat Jelek
0,20 - 0,39	Jelek
0,40 - 0,59	Cukup
0,60 - 0,79	Baik
0,80 - 1,00	Sangat Baik

6. Kemampuan Peserta Didik

Hasil penelitian untuk setiap peserta tes dikategorikan menjadi beberapa tingkat kemampuan (Azwar, 2010 : 148). Sebagai berikut

Tabel 7. Interval Kemampuan Berpikir Kreatif Peserta Didik

Rumus	Kategori
$M_i + 1,5 \times sb_i < \theta$	Sangat Tinggi
$M_i + 0,5 \times sb_i < \theta \leq M_i + 1,5 \times sb_i$	Tinggi
$M_i - 0,5 \times sb_i < \theta \leq M_i + 0,5 \times sb_i$	Sedang
$M_i - 1,5 \times sb_i < \theta \leq M_i - 0,5 \times sb_i$	Rendah
$0 < M_i - 1,5 \times sb_i$	Sangat Rendah

$$M_i = \frac{1}{2} (\Theta \text{ max} + \Theta \text{ min})$$

$$sb_i (\text{simpangan baku}) = \frac{1}{6} (\Theta \text{ maximum} - \Theta \text{ minimum})$$

Θ = kemampuan peserta tes, diambil dari hasil uji terbatas.

d. Uji Efektivitas CATIVE

Uji efektivitas dilakukan dengan pemberian angket kepada peserta didik untuk melihat respon pasca mengerjakan tes. Purwanto (2010) menyatakan angket dihitung dengan menghitung skor keseluruhan dengan rumus:

$$N = \frac{k}{Nk} \times 100\% \quad (3)$$

Keterangan:

N = Persentase kelayakan aspek

k = Skor hasil pengumpulan data

Nk = Skor keseluruhan nilai tertinggi

Hasil yang diperoleh kemudian dibandingkan dengan interval pada Tabel 8 sebagai berikut

Tabel 8. Kriteria Respon dan Konversi Efektifitas

Interval Persentase N	Konversi
$85\% < N \leq 100\%$	Sangat efektif
$72\% < N \leq 85\%$	Efektif
$58\% < N \leq 72\%$	Cukup efektif
$44\% < N \leq 58\%$	Tidak efektif
$N \leq 44\%$	Sangat tidak efektif