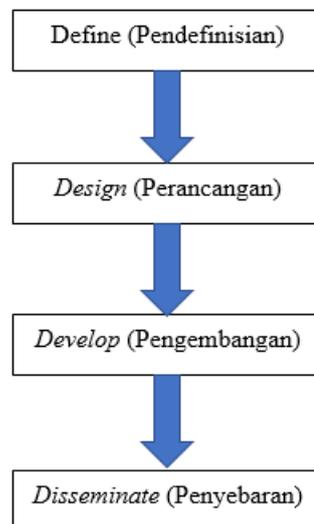


BAB III METODE PENELITIAN

3.1 Metode dan Desain Penelitian

Metode yang dilakukan dalam penelitian ini adalah RnD. RnD merupakan kependekan dari *Research and Development* (Penelitian dan Pengembangan) yaitu metode penelitian untuk mengembangkan dan menguji produk yang nantinya akan dikembangkan dalam dunia pendidikan. Terdapat berbagai macam model penelitian yang dapat dijadikan sebagai acuan dalam penelitian *Research and Development* (Amali dkk., 2019). Adapun model R&D yang digunakan dalam penelitian ini adalah model pengembangan 4D yang langkah-langkahnya ditampilkan dalam gambar 3.1.

Menurut (Thiagarajan dkk., 1974), model 4D terdiri dari empat tahap pengembangan. Tahap pertama *Define* atau sering disebut sebagai tahap analisis kebutuhan, tahap kedua adalah *Design* yaitu menyiapkan kerangka konseptual model, lalu tahap ketiga *Develop*, yaitu tahap pengembangan melibatkan uji validasi instrumen, dan terakhir adalah tahap *Disseminate*, yaitu implementasi pada sasaran sesungguhnya yaitu subjek penelitian.



Gambar 3.1 Langkah-langkah model pengembangan 4D

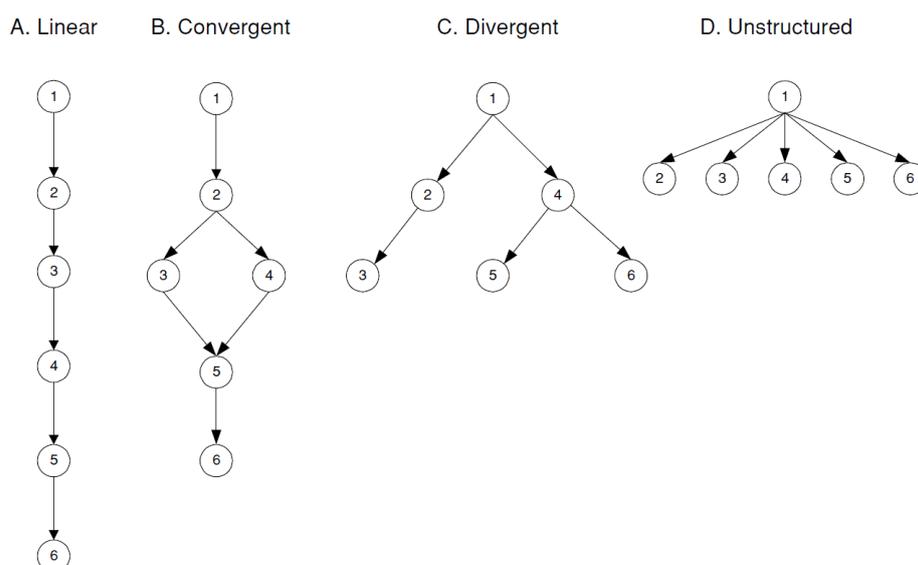
Adapun rincian tahapan pengembangan sebagai berikut:

1. Tahap *Define* (Pendefinisian)

Tahap pendefinisian ini bertujuan untuk mengidentifikasi apakah pengembangan instrumen OMUCHODA diperlukan, serta menetapkan tujuan pembelajaran yang jelas dan relevan pada materi dinamika gerak. Kegiatan ini diawali dengan analisis awal untuk menilai perlunya pengembangan instrumen, yang melibatkan studi literatur terkait instrumen dinamika gerak yang telah ada. Selanjutnya, dilakukan analisis tugas yang berfokus pada pemetaan tugas-tugas pembelajaran guna menetapkan tujuan pembelajaran. Hasil dari proses ini adalah perumusan tujuan pembelajaran yang terukur dan spesifik, yang akan menjadi dasar dalam pengembangan instrumen OMUCHODA.

2. Tahap *Design* (Perancangan)

Tahap perancangan bertujuan untuk merancang item-item instrumen OMUCHODA berdasarkan tujuan pembelajaran yang sudah dihasilkan pada tahap pendefinisian. Kegiatan pada tahap ini dimulai dengan menjabarkan tujuan pembelajaran menjadi sub tujuan pembelajaran atau sub atribut dan disusun dalam bentuk beberapa kombinasi dimulai dengan membuat atribut. Setelah atribut-atribut yang relevan diidentifikasi, langkah berikutnya adalah menyusunnya dalam bentuk hirarki. Adapun kemungkinan hirarki ditunjukkan pada gambar 3.1 yang terdiri dari tipe *linear*, *convergent*, *divergent*, atau *unstructured*.



Gambar 3.2 Tipe hirarki atribut

Setelah atribut-atribut dan hirarki tersusun dengan baik, langkah selanjutnya adalah menciptakan item-item yang akan merepresentasikan atribut tersebut. Hasil dari tahap ini adalah terbentuknya item-item instrumen OMUCHODA yang terstruktur berdasarkan atribut dan hirarki yang telah disusun, sehingga instrumen tersebut siap untuk dilanjutkan ke tahap selanjutnya untuk dinilai oleh ahli dan diujicobakan secara luas.

3. Tahap *Development* (Pengembangan)

Tahapan ini bertujuan untuk memastikan kualitas instrumen OMUCHODA melalui validasi ahli dan ujicoba produk. Kegiatan yang dilakukan pada tahapan ini yaitu sebagai berikut:

a. Validasi Ahli (*expert appraisal*)

Validasi yang dimaksud yaitu validasi isi yang terkait dengan konten. Validasi isi dilakukan berdasarkan pertimbangan profesional oleh kelompok pakar untuk menentukan kesahihan isi butir baik dari konten materi, konstruksi soal, maupun dari segi kejelasan bahasa. Dalam proses validasi ini para ahli diminta untuk mengoreksi dan memberi pertimbangan cakupan isi yang hendak diukur pada setiap item yang telah dibuat. Validasi ini akan dilakukan oleh 7 ahli atau dosen yang berkompeten dalam bidangnya.

a. Uji coba produk (*development testing*)

Setelah dilakukan validasi ahli kemudian dilakukan uji coba lapangan untuk mengetahui kualitas instrumen ditinjau dari segi tingkat kesukaran soal, nilai separasi instrumen, analisis fungsi distraktor, validitas dan reliabilitas instrumen, serta indeks konsistensi hirarki dari model yang sudah dibuat.

Hasil dari tahapan ini adalah instrumen OMUCHODA yang valid dan reliabel, sehingga layak digunakan sebagai perangkat evaluasi fisika, khususnya pada materi dinamika gerak.

3.2 Populasi dan sampel

Penelitian ini dilaksanakan di beberapa Sekolah Menengah Atas (SMA) yang ada di Indonesia. Penentuan sampel untuk uji coba dan implementasi

instrumen tes menggunakan teknik *sampling purposive*. Menurut Sugiyono (2014:124) teknik ini digunakan dengan pertimbangan tertentu sesuai dengan tujuan penelitian.

Subjek penelitian ini adalah peserta didik Kelas XI SMA yang telah mendapatkan materi dinamika gerak. Pemilihan kelas XI didasarkan pada asumsi bahwa pada tingkat ini, peserta didik sudah mempelajari materi dinamika gerak melalui proses pembelajaran yang mencakup teori dan aplikasi dinamika gerak dalam kehidupan sehari-hari. Adapun pemilihan lokasi sampel penelitian dilakukan dengan dua cara. Pertama, peneliti menyebarkan undangan ke beberapa sekolah yang ada di Bandung dan Batam dan memilih sekolah yang bersedia untuk dijadikan responden. Kedua, peneliti mengadakan lomba fisika khususnya untuk materi dinamika gerak secara nasional yang diikuti oleh seluruh sekolah yang tersebar di seluruh Indonesia. Berdasarkan metode pemilihan sampel tersebut, didapat jumlah subjek yang cukup besar, yaitu 188 peserta didik. Dengan jumlah ini diharapkan dapat memberikan data yang representatif dan memungkinkan analisis statistik yang lebih kuat dalam mengevaluasi pemahaman dan penerapan konsep-konsep fisika tersebut. Dengan demikian, hasil penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai evaluasi tingkat pemahaman peserta didik terhadap materi dinamika gerak.

3.3 Instrumen Penelitian

3.3.1 Lembar Validasi Ahli Konten

Untuk memastikan validitas konten instrumen yang telah dikembangkan, sebuah langkah penting yang diambil adalah menggunakan lembar validasi oleh para ahli. Lembar validasi oleh para ahli merupakan alat yang digunakan untuk mendapatkan masukan dan penilaian dari individu yang memiliki keahlian dan pengalaman dalam bidang terkait. Lembar validasi memberikan kerangka kerja yang sistematis untuk menilai sejauh mana instrumen dibuat sesuai dengan tujuan dan standar yang telah ditetapkan. Para ahli akan mengevaluasi aspek-aspek dalam validitas sebuah instrumen seperti Materi, Konstruksi, dan Bahasa.

3.3.2 Instrumen Tes OMUCHODA Dinamika Gerak

Penelitian ini menggunakan instrumen tes yang telah dikonstruksi melalui beberapa tahapan. Soal-soal yang digunakan pada penelitian ini yaitu dalam bentuk

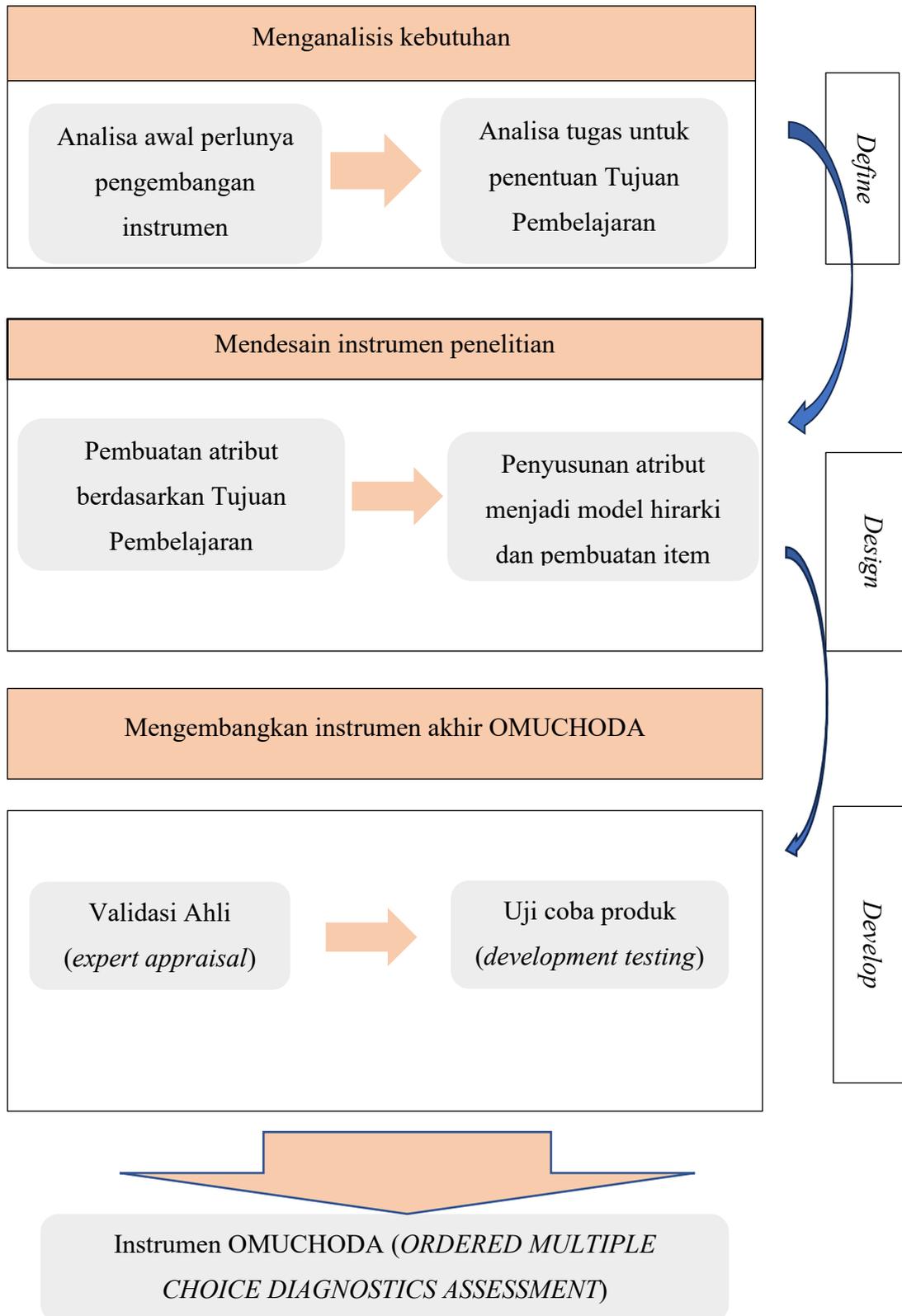
pilihan ganda. Jumlah soal yang disediakan disesuaikan berdasarkan analisis jumlah item pada tahap pembuatan instrumen. Instrumen tes yang diujikan telah divalidasi terlebih dahulu oleh ahli yang terdiri dari 8 orang ahli/pakar konten fisika. Instrumen tes dilengkapi dengan rubrik penilaian yang akan memudahkan dalam memberikan skor terhadap hasil jawaban siswa.

3.4. Prosedur Penelitian

Prosedur penelitian secara keseluruhan digambarkan pada diagram alur yang ditunjukkan melalui gambar 3.3. Berdasarkan gambar 3.3, prosedur penelitian terdiri dari tiga tahapan utama, yaitu menganalisis kebutuhan, mendesain instrumen penelitian, dan mengembangkan instrumen akhir OMOUCHODA (*Ordered Multiple Choice Diagnostics Assessment*). Pada tahap pertama, dilakukan analisis awal untuk menentukan perlunya pengembangan instrumen baru. Hal ini melibatkan identifikasi kebutuhan dan analisis tugas yang membantu menetapkan tujuan pembelajaran yang akan menjadi dasar pengembangan instrumen.

Tahap kedua adalah mendesain instrumen penelitian. Dalam tahap ini, atribut-atribut spesifik dibuat berdasarkan tujuan pembelajaran yang telah ditentukan sebelumnya. Atribut-atribut ini kemudian diorganisasikan menjadi model hirarki yang memetakan hubungan antara berbagai elemen, dan item-item untuk instrumen disusun berdasarkan model ini.

Tahap ketiga adalah pengembangan instrumen akhir OMOUCHODA. Setelah instrumen didesain, dilakukan validasi oleh para ahli (*expert appraisal*) untuk memastikan bahwa instrumen memenuhi standar keilmuan. Setelah validasi, instrumen diujicobakan melalui *development testing* untuk menguji kualitas instrumen seperti validitas konten, validitas item, reliabilitas instrumen, indeks konsistensi hierarki, tingkat kesukaran soal, nilai separasi instrumen, dan analisis fungsi distraktor. Hasil akhir dari proses ini adalah instrumen OMOUCHODA yang telah divalidasi dan siap digunakan untuk keperluan penelitian. Pendekatan penelitian ini melibatkan siklus yang menggabungkan proses *Define, Design, dan Develop* secara iteratif guna memastikan kualitas instrumen yang optimal.



Gambar 3.3 Prosedur penelitian

3.5 Teknik Pengolahan Data

Teknik pengolahan data terdiri atas :

3.4.1 Analisis Data Validitas Instrumen

Dalam proses validitas ini para ahli diminta untuk mengoreksi dan memberi pertimbangan cakupan isi yang hendak diukur pada setiap item yang telah dibuat. Validitas ini akan dilakukan oleh 7 ahli atau dosen yang berkompeten dalam bidangnya. Salah satu teknik yang digunakan untuk menghitung nilai validitas adalah metode *Content Validity Index* (CVI). CVI merupakan salah satu pendekatan dalam pengujian validitas isi. CVI menguji validitas isi dengan melibatkan tim ahli untuk menentukan setiap item dalam angket sesuai dengan konstruksinya. Dulu pendekatan CVI merupakan pendekatan validitas yang banyak digunakan dibidang keperawatan (Polit & Beck, 2006). Sekarang ini, pendekatan CVI banyak digunakan sebagai cara validasi di dunia pendidikan yang merupakan tempat untuk mengubah social melalui rekayasa sosial (Syamsudin, 2020). Dalam proses tersebut terdapat kombinasi antara manusia dan alam sekitar (Savira dkk., 2018). Pengembangan dalam pendekatan ini terus dilanjutkan dan kemudian penentuan CVI dibagi kedalam dua jenis, pertama yaitu validitas isi item individual (i-CVI) dengan persamaan sebagai berikut:

$$I-CVI = \frac{N_c}{n} \quad (11)$$

dengan:

N_c : Jumlah ahli yang menyatakan valid untuk setiap item

n : Jumlah keseluruhan ahli

dan kedua yaitu validitas isi secara keseluruhan (s-CVI) dengan persamaan sebagai berikut (Lynn, 1986):

$$S-CVI = \frac{\sum(I-CVI)}{p} \quad (12)$$

dengan:

p : jumlah item

Dalam penentuan apakah soal valid atau tidak ada beberapa kriteria yang bisa menjadi rujukan seperti ditampilkan pada tabel 3.1

Tabel 3.1
Kriteria penentuan validitas konten

Jumlah Ahli	Nilai CVI yang dapat diterima	Sumber Rekomendasi
2 ahli	Minimal 0,80	Davis (1992)
3 sampai 5 ahli	Harus 1,00	Polit & Beck (2006), Polite et,al (2007)
Minimal 6 ahli	Minimal 0,83	Polit & Beck (2006), Polite et,al (2007)
6 sampai 8 ahli	Minimal 0,83	Lynn (1986)
Minimal 9 ahli	Minimal 0,78	Lynn (1986)

3.4.2 Model-Data Fit

Kesesuaian model data mengungkapkan konsistensi respons antara pola yang diharapkan dalam hirarki dan pola yang diamati yang dihasilkan oleh siswa dalam merespons item diagnostik. Indeks Konsistensi Hierarki/*Hierarchy Consistency Index* (HCI) dapat digunakan untuk mengevaluasi kesesuaian ini (Cui & Leighton, 2009). Adapun nilai HCI dicari dengan menggunakan persamaan berikut:

$$HCI_i = 1 - \frac{2 \sum_{j \in S_{correct_i}} \sum_{j \in S_j} X_{ij} (1 - X_{ig})}{N_{ci}} \quad (13)$$

dimana:

- $S_{correct_i}$: item yang dijawab benar oleh responden
- X_{ij} : skor responden ke-i untuk item j (1 atau 0)
- S_j : item-item dimana atributnya diukur oleh item j
- X_{ig} : skor responden ke-i untuk item g dimana berhubungan dengan S_j
- N_{ci} : total perbandingan item yang dijawab benar oleh responden i

HCI menghasilkan indeks yang berkisar dari ketidaksesuaian sempurna sebesar -1 hingga kesesuaian sempurna sebesar 1. Cui (2007) menyatakan bahwa nilai-nilai HCI di atas 0,60 mengindikasikan kesesuaian model-data yang sedang, sementara nilai di atas 0,80 menunjukkan kesesuaian yang sangat baik.

3.4.3 Validitas Butir

Validitas butir dalam penelitian ini dianalisis menggunakan Rasch model. Menurut Boone, dkk. (2014), kriteria yang digunakan kriteria yang digunakan untuk memeriksa butir soal yang sesuai adalah

1. Nilai Outfit Mean Square (MNSQ) yang diterima : $0,5 < MNSQ < 1,5$

2. Nilai outfit Z-standard (ZSTD) yang diterima: $-2,0 < ZSTD < +2,0$

Pertama dilihat dulu apakah MNSQ yang dalam interval. Jika masih dalam interval yang wajar maka soal tergolong valid dan tidak perlu melihat nilai ZSTD. Akan tetapi jika nilai MNSQ tidak bagus, asalkan nilai ZSTD nya masih dalam interval wajar, soal masih bisa dikatakan valid.

3.4.4 Reliabilitas Instrumen

Pada penelitian ini analisis reliabilitas menggunakan Cronbach's Alpha (α) untuk menilai instrumen keseluruhan secara klasik dan *Rasch item reliability* untuk menilai reliabilitas pada *Rasch model*. Kriteria Cronbach's Alpha ditampilkan pada tabel 3.2 berikut:

Tabel 3.2
Kriteria nilai Cronbach's Alpha (Reliabilitas)

Nilai Cronbach's Alpha (Reliabilitas)	Kriteria
$\alpha < 0,50$	Buruk
$0,50 \leq \alpha < 0,60$	Jelek
$0,60 \leq \alpha < 0,70$	Cukup
$0,70 \leq \alpha \leq 0,80$	Bagus
$\alpha > 0,80$	Bagus sekali

Adapun kriteria yang digunakan untuk menentukan reliabilitas instrumen Rasch ditampilkan pada tabel 3.3 berikut:

Tabel 3.3
Kriteria nilai reliabilitas instrumen Rasch

Nilai reliabilitas instrumen	Kriteria
$r < 0,67$	Lemah
$0,67 \leq r < 0,80$	Cukup
$0,80 \leq r < 0,91$	Bagus
$0,91 \leq r \leq 0,94$	Bagus sekali
$r > 0,94$	Istimewa

3.4.5 Tingkat Kesukaran Soal

Untuk mengetahui tingkat kesukaran butir soal dalam Rasch model maka dilihat nilai logit tiap butir soal yang dapat dilihat pada kolom *measure (im)*. Nilai logit yang tinggi menunjukkan tingkat kesulitan soal yang tinggi. Tabel 3.4 menunjukkan kategori kelompok soal berdasarkan tingkat kesulitannya. Nilai Standar Deviasi (SD) dalam penelitian ini adalah 0,93.

Tabel 3.4
Kategori tingkat kelompok soal berdasarkan tingkat kesulitannya

Nilai tingkat kesukaran butir	Kriteria
$im > +0,93$	Sangat sulit
$0,0 \logit \leq im \leq +0,93$	Sulit
$-0,93 \leq im < 0,0$	Sedang
$im < -0,93$	Mudah

3.4.6 Nilai Separasi Instrumen

Pada Rasch model nilai separasi instrumen ditentukan dengan menggunakan cara mengidentifikasi kelompok responden berdasarkan indeks separasi responden. Semakin besar nilai separasi item maka kualitas instrumen dalam hal keseluruhan responden dan butir soal makin bagus, karena dapat mengidentifikasi kelompok responden dan kelompok butir (Sumintono & Widhiarso, 2015). Persamaan lain untuk mengetahui pengelompokan secara lebih teliti digunakan persamaan strata seperti ditunjukkan pada persamaan berikut:

$$H = \frac{[(4 \times \text{SEPARATION}) + 1]}{3} \quad (14)$$

3.4.7 Analisis Fungsi Distraktor

Dalam analisis Rasch model, distraktor dalam soal dianggap berfungsi dengan baik apabila nilai *mean ability* naik secara konsisten. Kenaikan yang konsisten ini menandakan bahwa distraktor tersebut berhasil membedakan antara siswa dengan kemampuan yang berbeda secara efektif. Dengan kata lain, siswa dengan kemampuan lebih rendah cenderung memilih jawaban yang salah lebih sering daripada siswa dengan kemampuan lebih tinggi.

Sebaliknya, distraktor yang tidak berfungsi dengan baik ditandai dengan nilai *mean ability* yang tidak menunjukkan kenaikan yang konsisten. Perubahan yang tidak konsisten ini menunjukkan bahwa distraktor tersebut tidak berhasil membedakan antara siswa dengan kemampuan yang berbeda dengan baik. Distraktor semacam ini tidak memberikan informasi yang jelas tentang perbedaan kemampuan siswa, yang mengurangi efektivitas soal tersebut dalam menilai kemampuan siswa secara akurat.