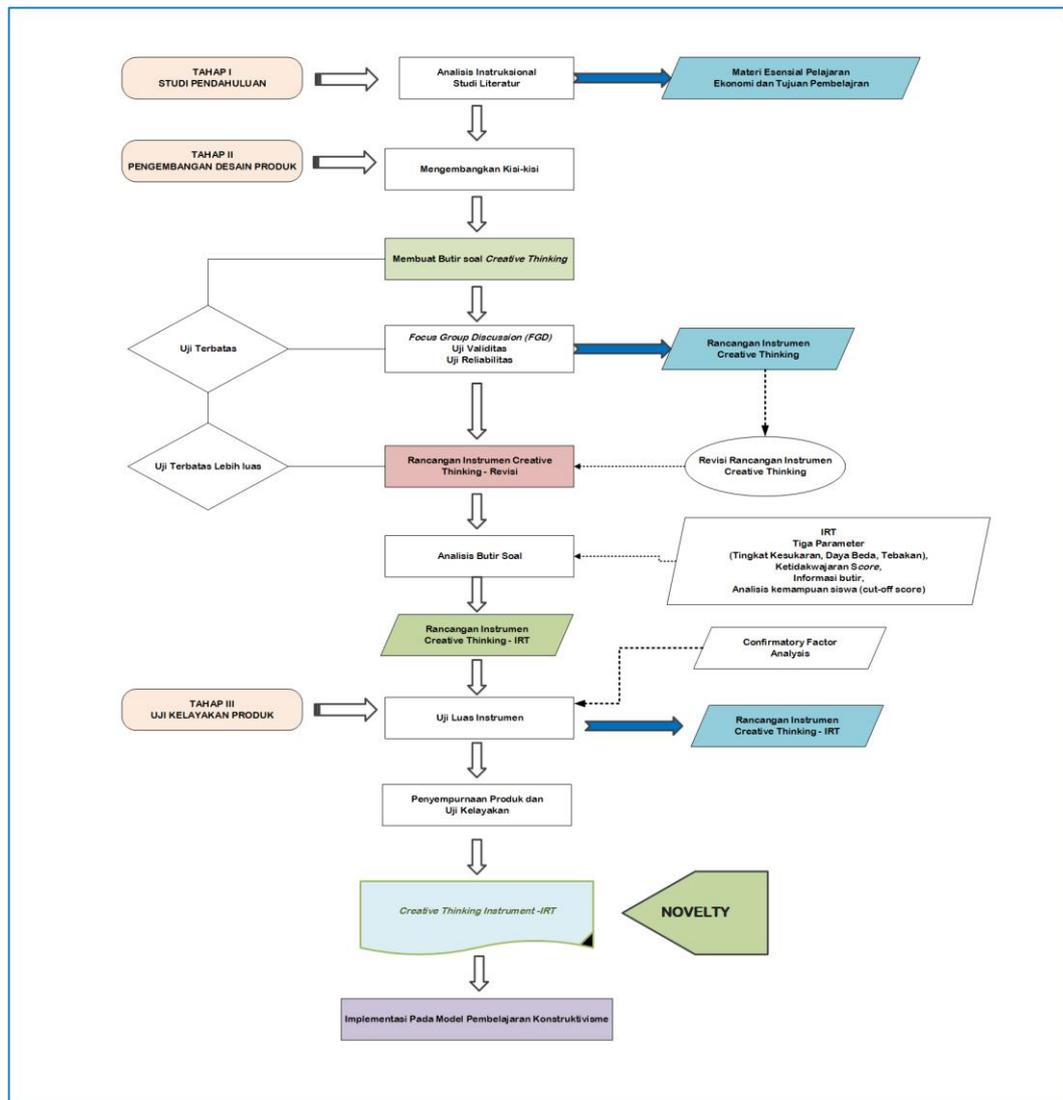


## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Metode Penelitian**

*Research and development* (R&D) adalah proses yang digunakan untuk mengembangkan instrumen *creative thinking* dalam pembelajaran ekonomi. *Research and development* adalah proses berbasis industri di mana temuan penelitian diterapkan pada desain produk dan prosedur baru. Ini kemudian diuji secara menyeluruh di lapangan, dinilai, dan ditingkatkan hingga memenuhi standar yang telah ditentukan sebelumnya untuk kualitas, kemandirian, atau kriteria yang sebanding (Gall & Gall, 2007). Pendekatan penelitian ini dianggap cocok untuk mengembangkan instrumen *creative thinking* dalam pendidikan ekonomi. Model penelitian dan pengembangan yang dijelaskan oleh Gall & Gall (2007) terdiri dari sepuluh tahap berikut: 1) *preliminary study* (pengumpulan informasi dan penelitian); 2) *planning research* (perencanaan); 3) *design development* (pengembangan produk awal); 4) *preliminary product*; 5) *revision of limited field trial results* (revisi produk utama); 6) *main field test*; 7) *revision of wider field test results* (revisi produk operasional); 8) *feasibility test* (pengujian lapangan operasional); 9) *final revision of the feasibility test results* (revisi produk akhir); 10) *dissemination and implementation of the final product* (diseminasi dan implementasi). Gambar 3.1 menunjukkan proses implementasi R&D secara grafis.



**Gambar 3. 1** Prosedur Pelaksanaan R&D Pengembangan Instrumen *Creative Thinking* Dalam Pembelajaran Ekonomi

### 3.2 Responden

Penelitian ini dilakukan pada siswa kelas X di SMA Negeri Kota Tasikmalaya dan Kota Bandung tahun akademik 2023-2024. Pemilihan sekolah didasarkan pada kesesuaian dengan kurikulum yang digunakan, yaitu Kurikulum 2013. Untuk memenuhi persyaratan *Item Response Theory* (IRT), instrumen ini harus diujikan pada jumlah responden yang besar, yaitu antara 200 hingga 1000 orang, karena semakin banyak responden yang terlibat, semakin baik hasil analisis IRT yang diperoleh (Susetyo, 2015). Dalam penelitian ini, responden terdiri dari

Ai Nur Solihat, 2024

**PENGEMBANGAN INSTRUMEN CREATIVE THINKING DENGAN MENGGUNAKAN ITEM RESPONSE THEORY (IRT) DALAM PEMBELAJARAN EKONOMI MODEL KONSTRUKTIVISME**  
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

906 siswa kelas X SMA Negeri di Kota Tasikmalaya dan Kota Bandung. Pertimbangan dalam pemilihan responden didasarkan pada kebutuhan akan tingkat kemampuan yang heterogen. Responden dipilih secara acak dari SMAN 1 Kota Tasikmalaya, SMAN 5 Kota Tasikmalaya, SMAN 4 Kota Bandung, SMAN 8 Kota Bandung, dan SMAN 24 Kota Bandung. Proses pemilihan responden dilakukan sesuai dengan etika penelitian yang berlaku, dengan memastikan bahwa tidak ada subjek yang dirugikan dan bahwa semua partisipasi bersifat sukarela. Rincian jumlah responden pada setiap sekolah dapat dilihat pada Tabel 3.1

**Tabel 3. 1 Rincian Responden Berdasarkan Sekolah dan Jenis Kelamin**

No	Sekolah	Jenis Kelamin		Total
		Laki-laki	Perempuan	
	SMAN 4 Kota Bandung	33	36	69
	SMAN 8 Kota Bandung	62	80	142
	SMAN 24 Kota Bandung	33	28	61
	SMAN 1 Kota Tasikmalaya	179	285	464
	SMAN 5 Kota Tasikmalaya	82	88	170
	<b>Total</b>	389	517	<b>906</b>

### 3.3 Sampel

Sampel dalam penelitian ini merupakan sample butir dari instrumen *creative thinking* sebanyak 60 butir *multiple choice* yang mengacu pada dimensi *Torrance Test Creative Thinking* yaitu *fluency*, *flexibility*, *originality* dan *elaboration*. Setiap item soal diimplementasikan dalam materi pelajaran ekonomi. Pemilihan materi berdasarkan hasil analisis karakteristik materi ekonomi yang sesuai dengan dengan karakteristik kemampuan *creative thinking*. Adapun rincian jumlah soal pada setiap dimensi dapat di lihat pada Tabel 3.2.

**Tabel 3. 2 Rincian jumlah soal untuk setiap dimensi *creative thinking*.**

No	Dimensi <i>Creative Thinking</i>	No Soal	Jumlah
1.	<b>Berpikir lancar (<i>fluency</i>):</b> Menemukan banyak solusi terhadap suatu masalah dan memilih satu di antaranya adalah ciri ide atau jawaban yang produktif.	1, 2, 3, 4, 6, 9, 12, 13, 24, 33, 42.	11

No	Dimensi <i>Creative Thinking</i>	No Soal	Jumlah
2	<b>Berpikir luwes (<i>fleksibel</i>):</b> Kemampuan untuk mengubah strategi dan pendekatan, berpikir dengan cara baru, adalah yang membuat seseorang mahir dalam menghasilkan konsep yang konsisten.	7, 15, 25, 26, 31, 35, 38, 45, 48, 50, 51, 57, 58, 59.	14
3	<b>Berpikir orisinal (<i>originality</i>):</b> memberikan respon yang berbeda dari yang lain dan tidak biasa dari apa yang diberikan kebanyakan orang ditandai oleh keterampilan memecahkan masalah.	10, 11, 16, 17, 18, 19, 20, 21, 29, 30, 32, 34, 36, 37, 40, 41, 44, 46, 54, 56, 60.	21
4	<b>Berpikir terperinci (<i>elaboration</i>):</b> mengembangkan suatu ide ditandai dengan kemampuan untuk masuk ke detail dalam memecahkan suatu masalah.	5, 8, 14, 22, 23, 27, 28, 39, 43, 47, 49, 52, 53, 55.	14
<b>TOTAL</b>			<b>60</b>

### 3.4 Instrumen Penelitian

Instrumen merupakan alat bantu yang dipilih dan digunakan oleh peneliti dalam kegiatan pengumpulan data. Instrumen yang digunakan dalam penelitian ini disesuaikan dengan rumusan masalah yang akan dipecahkan, yaitu:

#### 1. Angket (*Questionnaire*)

Angket atau kuesioner merupakan suatu teknik atau cara pengumpulan data secara tidak langsung yang berisi sejumlah pertanyaan yang harus dijawab oleh responden. Dalam penelitian ini angket digunakan pada tahap analisis kebutuhan. Sejumlah pernyataan diberikan kepada 24 orang guru mata pelajaran ekonomi SMA/MA negeri di Kota Tasikmalaya untuk mengetahui seberapa perlu instrumen *creative thinking* dalam pembelajaran ekonomi dengan menggunakan *Item Response Theory* (IRT).

## 2. Wawancara

Wawancara merupakan salah satu bentuk teknik pengumpulan data yang banyak digunakan dalam penelitian dalam bentuk pertanyaan lisan kepada responden secara terstruktur ataupun tidak terstruktur. Dalam penelitian ini wawancara dilakukan pada tahap analisis kebutuhan sebagai bentuk konfirmasi atas jawaban yang diberikan responden pada angket (*questionnaire*).

## 3. Tes

Tes merupakan serangkaian pertanyaan atau latihan yang digunakan untuk mengukur kemampuan tertentu. Dalam penelitian ini tes yang digunakan adalah tes kemampuan berpikir kreatif (*creative thinking*). Tes ini dilakukan pada tahap uji coba terbatas dan uji coba luas.

### 3.5 Prosedur Penelitian

Dengan mengadopsi dan modifikasi model Borg and Gall (2007), maka prosedur pengembangan instrument pengukuran *creative thinking* dalam pembelajaran ekonomim sebagai berikut.

#### 3.5.1 Tahap 1 Perencanaan Produk

##### 3.5.1.1 Studi Pendahuluan (Research and Information Collecting)

###### a. Analisis kebutuhan

Analisis kebutuhan dilakukan untuk mengetahui apakah produk yang akan dikembangkan merupakan suatu hal yang penting bagi kemajuan Pendidikan, untuk menentukan estimasi waktu yang dibutuhkan dalam pengembangan produk tersebut. Analisis kebutuhan dapat dilakukan dengan survey, dan wawancara langsung ke lapangan. Hasil analisis kebutuhan dapat digunakan sebagai dasar pertimbangan untuk mengembangkan instrument pengukuran *creative thinking* dalam pembelajaran ekonomi.

###### b. Studi Pustaka/studi literatur

Studi Pustaka/studi literatur dilakukan sebagai upaya untuk mengenali dan memahami produk yang akan dikembangkan melalui buku-buku referensi, temuan hasil riset, dan informasi lain terkait produk yang akan dikembangkan. Pada tahapan ini, dilakukan analisis terhadap standar isi mata pelajaran ekonomi.

### 3.5.1.2 Perencanaan

Pada tahap ini, peneliti merumuskan tujuan pengembangan model, sasaran dan penggunaan model, deskripsi mengenai tahapan-tahapan penggunaan model. Kegiatan yang dilakukan pada tahap perencanaan ini adalah sebagai berikut.

- a. Merumuskan tujuan pengembangan model
- b. Menentukan materi pembelajaran yang akan dikembangkan menjadi instrument pengukuran *creative thinking*. Berdasarkan hasil studi pendahuluan yang telah dilakukan, materi yang dikembangkan menjadi instrumen *creative thinking* adalah Masalah Ekonomi Dalam Sistem Perekonomian, Peran dan Pelaku Ekonomi Dalam Kegiatan Ekonomi, Keseimbangan Pasar dan Struktur Pasar.
- c. Melakukan analisis instruksional kesesuaian antara kompetensi dasar, indikator pengukuran *creative thinking* dan indicator butir soal yang akan dikembangkan.

### 3.5.2 Tahap II Pengembangan Produk

#### 3.5.2.1 Pengembangan Desain (*Develop Preliminary of Product*)

Kegiatan yang dilakukan pada tahap ini adalah mengembangkan kisi-kisi instrument pengukuran *creative thinking*.

#### 3.5.2.2 Uji Terbatas (*Preliminary Field Testing*)

Pada tahap ini dilakukan uji coba terbatas yaitu melakukan uji coba instrumen *creative thinking* yang dikembangkan meliputi uji coba desain produk, uji validitas dan reliabilitas instrument, dan melakukan FGD. Tahap ini bersifat terbatas, baik substansi maupun pihak-pihak yang terlibat.

#### A. *Focus Group Discussion* (FGD)

*Focus Group Discussion* (FGD) dibagi menjadi dua tempat di SMAN 5 Kota Tasikmalaya, dan SMAN 4 Kota Bandung yang dihadiri oleh guru ekonomi yang tergabung dalam Musyawarah Guru Mata Pelajaran (MGMP) dan juga tim penyusun soal OSN Mata Pelajaran Ekonomi. Untuk pelaksanaan FGD di SMAN 5 Kota Tasikmalaya FGD dilaksanakan pada tanggal 30 Agustus 2023 dihadiri oleh 6 orang guru ekonomi sedangkan di Kota Bandung dihadiri oleh 3 orang guru

ekonomi dilaksanakan pada tanggal 3 September 2023. Yang menjadi fokus pembahasan dalam pelaksanaan FGD ini terkait validitas isi, telaah soal menyangkut konsistensi terkait konstruk, materi, dan bahasa.

## **B. Uji Persyaratan Tes**

### **1) Validitas isi dengan formula uji kecocokan**

Validitas isi merupakan validitas yang akan mengecek kecocokan di antara butir-butir tes yang dibuat dengan indikator, materi atau tujuan pembelajaran yang telah ditetapkan. Butir tes yang dinyatakan valid, jika butir-butir yang dibuat secara tepat mengukur indikator. Suatu tes dikatakan memiliki validitas isi, apabila butir-butir yang disusun sesuai dengan materi pelajaran dan indikator yang telah ditetapkan. Validitas isi lebih cocok digunakan dalam bidang pendidikan atau pembelajaran di sekolah.

Untuk mengetahui validitas isi pada hasil belajar, perlu dilakukan pencocokan antara materi dan butir tes dengan jalan mencocokkan antara isi butir tes hasil belajar dan indikator-indikator yang telah ditetapkan pada setiap topik pembelajaran. Suatu tes telah memenuhi validitas guru/dosen ahli. Butir tes dinyatakan valid jika kecocokannya dengan indikator mencapai lebih dari 50% (Susetyo, 2015). Rumus yang digunakan adalah:

$$\text{Persentase} = \frac{f}{\sum f} \times 100\%$$

Di mana:

$f$  = frekuensi cocok menurut penilai

$\sum f$  = jumlah penilai.

### **2) Validitas isi dengan Formula Aiken**

Pada umumnya, validitas isi yang dilakukan oleh beberapa pakar (*expert judgement*) dinyatakan dalam bentuk dikotomi (ya -tidak, setuju-tidak setuju, valid-tidak valid, cocok-tidak cocok) atau peringkat (validitas tinggi-validitas sedang-validitas rendah) (Aiken, 1980). Validitas isi bukan hanya meringkas pernyataan dari consensus beberapa pakar, tetapi harus dipastikan bahwa pengukuran validitas

isi benar mengukur konsep yang akan diukur. Oleh karena itu, dalam penelitian ini juga dihitung indeks validitas Aiken dengan menggunakan rumus:

$$V = \sum_{i=1}^{c-1} \frac{in_1}{N(c-1)}$$

Dimana:

$i = (r+1)$  sampai  $(r + t + 1)$

$n_i$  = banyaknya pilihan pakar pada nilai  $I$

$N = \sum n_i$

$c$  = nilai pagu pada pilihan validitas (misal 5)

$r$  = nilai floor pada pilihan validitas (misal 1)

Dalam menentukan signifikansi statistik nilai indeks validitas Aiken, maka nilai indeks validitas tersebut dibandingkan dengan nilai table probabilitas nilai  $V$  untuk 2 sampai 7 kategori rating ( $c$ ) dan 2 pilihan untuk 25 item ( $m$ ) atau penilai ( $n$ ). Dua pilihan tersebut berdasarkan probabilitas 0,01 atau 0,05 (Aiken, 1985).

### 3) Uji Reliabilitas

Tahap uji terbatas dalam penelitian ini yaitu menghitung nilai reliabilitas instrumen *creative thinking*. Reliabilitas merupakan keandalan koefisien yang menunjukkan tingkat keajegan atau konsistensi hasil pengukuran suatu tes. Konsistensi berkaitan dengan tingkat kesalahan hasil suatu tes berupa skor. Perhitungan nilai reliabilitas menggunakan Kuder Richardson (KR) secara langsung pada seluruh butir tes. Perhitungan korelasi KR menghendaki kesetaraan pada semua butir tes dalam perangkat ukur. Alasan pemilihan perhitungan nilai reliabilitas menggunakan KR yaitu bahwa data yang dipergunakan dalam perhitungan KR berbentuk dikotomi dengan skor butir 1 untuk jawaban benar dan 0 untuk jawaban salah (Susetyo, 2015). Adapun rumus yang dipergunakan dalam perhitungan reliabilitas dengan menggunakan KR ini sebagai berikut.

$$\rho_{KR20} = \frac{k}{k-1} \left[ 1 - \frac{\sum pq}{\sigma_x^2} \right]$$

Keterangan:

$p$  = proporsi jawaban benar

$q$  = proporsi jawaban salah

$k$  = jumlah butir

$\sum pq$  = jumlah perkalian jawaban benar salah

$\rho_{KR20}$  = koefisien reliabilitas

$\sigma_X^2$  = varian skor tes

$N$  = jumlah responden

### 3.5.2.3 Revisi Hasil Uji Terbatas (*Main Product Revision*)

Langkah ini merupakan perbaikan model atau desain berdasarkan uji lapangan terbatas. Penyempurnaan produk awal akan dilakukan setelah dilakukan uji coba lapangan secara terbatas. Pada tahap penyempurnaan produk awal ini, lebih banyak dilakukan dengan pendekatan kualitatif. Evaluasi yang dilakukan lebih pada evaluasi proses sehingga perbaikan yang dilakukan bersifat perbaikan internal. Pada penelitian ini revisi hasil uji coba terbatas dilakukan dengan memperbaiki item-item soal berdasarkan konten, konstruk, dan bahasa sesuai dengan penilaian ahli pada proses *Focus Group Discussion* (FGD).

### 3.5.2.4 Uji Luas (*Main Field Test*)

Langkah uji produk secara lebih luas. Langkah ini meliputi uji efektivitas desain produk, uji efektivitas desain. Pada tahap ini item soal pada instrumen *creative thinking* di uji cobakan pada 453 orang responden. Hasil uji coba luas ini kemudian dilakukan analisis item soal dengan menggunakan *Item Response Theory* (IRT) dengan model logistik 3 parameter. Hasil uji produk lebih luas ini berupa desain produk yang efektif baik dilihat dari tingkat kesukaran, daya beda dan tebakan.

### 3.5.2.5 Revisi Hasil Uji Luas (*Operational Product Revision*)

Pada tahap ini dilakukan perbaikan-perbaikan item soal dari aspek konstruk, dan Bahasa berdasarkan hasil pada uji luas. Penyempurnaan produk dari hasil uji luas ini akan memantapkan produk yang kita kembangkan. Perbaikan pada tahap

ini bersifat internal, penyempurnaan produk ini didasarkan pada evaluasi hasil analisis butir soal dengan menggunakan IRT 3 parameter logistik.

### **3.5.3 Tahap III Uji Kelayakan Produk**

#### **1. Uji Kelayakan (*Operational Filed Testing*)**

Uji kelayakan dilakukan dalam skala besar. Pada tahap ini dilakukan uji efektivitas dan adaptabilitas desain produk. Uji efektivitas dan adaptabilitas desain melibatkan para calon pemakai produk, hasil uji lapangan adalah diperoleh model desain yang siap diterapkan, baik dari sisi substansi maupun metodologi. Pada tahap ini, instrumen *creative thinking* yang telah di uji pada tahap main field test, diujikan kembali kepada 906 orang responden, kemudian dilakukan analisis uji kelayakan dengan menggunakan *Confirmatory Factor Analysis* (CFA).

#### **2. Revisi Final Hasil Uji Kelayakan (*Final Product Revision*)**

Tahap ini dilakukan untuk penyempurnaan produk yang sedang dikembangkan. Penyempurnaan produk akhir dipandang perlu untuk lebih akuratnya produk yang dikembangkan. Jika terdapat item yang dinyatakan tidak layak pada tahap uji kelayakan pertama, maka item tersebut dihilangkan dari model pengukuran (trimming) kemudian item-item yang dinyatakan layak diestimasi ulang. Pada tahap ini sudah didapatkan suatu produk yang tingkat ini efektivitasnya dapat dipertanggungjawabkan. Hasil penyempurnaan produk akhir memiliki nilai generalisasi yang dapat diandalkan. Dari hasil uji kelayakan yang dilakukan diperoleh item-item soal yang dapat dinyatakan valid dan reliabel berdasarkan hasil analisis CFA.

#### **3. Implementasi dan Desiminasi Product Akhir (*Dissemination and Implementation*)**

##### **a) Implementasi**

Untuk mengetahui efektivitas dari instrumen yang sudah dinyatakan layak, maka instrumen test *creative thinking* ini di implementasikan pada model-model

pembelajaran konstruktivisme dengan menggunakan metode eksperimen. Model pembelajaran konstruktivisme yang digunakan yaitu *Discovery Learning*, *Group Investigation*, dan *Problem Solving*. Pemilihan model pembelajaran tersebut berdasarkan pada teori belajar konstruktivisme yang menyatakan bahwa fokus teori pada pembelajaran berpusat pada siswa dan peran aktif siswa dalam membangun pemahaman dapat menumbuhkan pemikiran kreatif.

Metode eksperimen yang digunakan pada tahap implementasi ini yaitu quasi-experiment dengan *The One-Group Pretest-Posttest Design*. Menurut Fraenkel dan Wallen (2012) menyebutkan bahwa : “*one-group pretest-posttest design, a single group is measured or observed not only after being exposed to a treatment of some sort, but also before This design is better than the one-shot case study (the researcher at least knows whether any change occurred), but it is still weak.*” Dalam menganalisis data, skor *pretest* masing-masing individu adalah dikurangi dari skor *posttest*-nya, sehingga memungkinkan analisis *gain* atau perubahan. Desain tersebut dapat digambarkan pada Tabel 3.3.

**Tabel 3. 3 *The One-Group Pretest-Posttest Design***

O <sub>1</sub>	X	O <sub>2</sub>
----------------	---	----------------

Keterangan :

- X : Penerapan model pembelajaran pada kelas eksperimen
- O<sub>1</sub> : *Pre Test* (Tes awal sebelum perlakuan) pada masing-masing kelompok
- O<sub>2</sub> : *Post test* (Tes akhir setelah perlakuan ) pada masing-masing kelompok

### 3.6 Teknik Analisis Data

#### 3.6.1 Analisis Informasi Butir Soal

*Item Response Theory* (IRT) memiliki persyaratan tertentu yang harus dipenuhi agar model dapat diterapkan secara akurat. Berikut adalah beberapa persyaratan utama dalam analisis IRT.

### 1) Unidimensi

Unidimensionalitas merupakan asumsi penting dalam Teori Respons Butir (IRT). Konsep ini berkaitan dengan gagasan bahwa tes atau alat yang diteliti hanya menilai satu dimensi atau konstruk kemampuan ( $\theta$ ), seperti kemampuan *creative thinking* dalam pembelajaran ekonomi. Unidimensionalitas sangat penting dalam model IRT karena model ini menunjukkan bahwa variasi dalam respon peserta terhadap butir-butir tes merupakan hasil dari perbedaan tingkat kemampuan pada satu dimensi yang dinilai oleh tes. Model IRT efektif ketika ada satu dimensi utama yang dianalisis. Dali (Susetyo, 2015) menyatakan bahwa teknis analisis yang digunakan dalam pemeriksaan unidimensi pada penelitian ini dengan menggunakan analisis factor, yaitu dengan melihat eigenvalue factor pertama lebih besar daripada factor kedua, dan eigenvalue kedua hamper sama dengan faktor berikutnya.

### 2) Invariansi kelompok

Invariansi parameter artinya karakteristik butir soal tidak tergantung pada distribusi parameter kemampuan peserta tes dan parameter yang menjadi ciri peserta tes tidak bergantung dari ciri butir soal. Kemampuan seseorang tidak akan berubah hanya karena mengerjakan tes yang berbeda pada tingkat kesulitannya dan parameter butir tes tidak akan hanya karena diujikan pada kelompok peserta tes yang berbeda tingkat kemampuannya. Hanlbleton et al. (1992) menyebutkan bahwa invariansi parameter kemampuan dapat diselidiki dengan mengajukan dua perangkat tes atau lebih yang memiliki tingkat kesukaran yang berbeda pada sekelompok peserta tes. Invariansi parameter kemampuan akan terbukti jika hasil estimasi kemampuan peserta tes tidak berbeda walaupun tes yang dikerjakan berbeda tingkat kesulitannya. Invariansi parameter butir dapat diselidiki dengan mengujikan tes pada kelompok peserta yang berbeda. Invariansi parameter butir terbukti jika hasil estimasi parameter butir tidak berbeda walaupun diujikan pada kelompok peserta yang berbeda tingkat kemampuannya. Pada penelitian ini analisis invariansi menunjukkan bahwa tidak terjadi invariansi parameter hasil estimasi (Tingkat kesukaran,

daya beda dan tebakan) pada kelompok laki-laki dan Perempuan. Adapun analisis insvariansi dilakukan dengan mengkorelasikan hasil estimasi parameter berdasarkan kelompok laki-laki dan Perempuan. Rumus korelasi yang digunakan yaitu korelasi product moment (Susetyo, 2015).

$$r = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

### 3) Uji Independensi Lokal

Persyaratan interdependensi local menghendaki pada lokasi yang sama dengan kemampuan responden yang sama, maka probabilitas menjawab benar  $P(\theta)$  untuk butir berbeda adalah independent satu butir tes terhadap butir tes yang lainnya. Adapun penghitungan independensi local pada penelitian ini melalui rumus probabilitas (Susetyo, 2015).

$$P(X_1 X_2 X_3 \dots X_N) = P(X_1)P(X_2)P(X_3) \dots P(X_N)$$

### 4) Model Logistik

Pemeriksaan model logistik dimaksudkan untuk melihat kecocokan data empiris pada setiap butir tes dengan model logistik. Teknik analisis yang digunakan untuk melihat kecocokan adalah  $\chi^2$ , dengan membandingkan parameter butir dengan kriteria pada tabel  $\chi^2$  atau menggunakan *scatterplot* atau grafik model logistik (Susetyo, 2015). Adapun rumus yang digunakan yaitu:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Keterangan:

$\chi^2$  = chi kuadrat

$O_i$  = frekuensi observasi

$E_i$  = frekuensi harapan (teori)

dk = k - 1

### 1) Analisis *Item Response Theory* (IRT)

Dalam penelitian ini analisis informasi butir menggunakan tiga parameter berupa tingkat kesukaran, daya pembeda dan tebakan. Berikut rumus probabilitas menjawab benar untuk model logistik:

$$P_i(\theta_j) = c_i + \frac{(1 - c_1)}{1 + e^{-D a_i(\theta_j - b_1)}}$$

Keterangan:

$P_i(\theta_j)$  = probabilitas peserta tes ke-j dengan taraf kemampuan ( $\theta_j$ ) menjawab butir ke-I dengan benar (butir  $I=1,2,3, \dots n$ )

$a_i$  = daya butir tes ke-i

$b_i$  = tingkat kesukaran butir tes ke -i

$e$  = besaran konstanta 2,718

$D$  = besaran konstanta 1,7 (1,702)

### 2) Analisis Ketidakwajaran Skor

Metode pemetaan butir (*item mapping*) digunakan untuk memperoleh informasi ketidakwajaran skor dalam penelitian ini. Retnawati (2014) menyatakan bahwa metode ini dilakukan dengan mengsetimasi log odds tingkat kesulitan yang merupakan parameter tingkat kesulitan terlebih dahulu. Dengan diestimasi tingkat kesulitan, sifat pengukuran yang pada intinya mengukur variabel laten, dengan diestimasi skala kemampuan ( $\theta$ ) dapat diketahui. Untuk dapat mengestimasi parameter kemampuan, diperlukan pola respons peserta didik terhadap tes. Ini menunjukkan bahwa meskipun pada proses penetapan *cut of score*, proporsi menjawab benar ditentukan oleh ahli, tetap diperlukan data empiris dari lapangan untuk estimasi parameter butir. Dengan menggunakan metode ini, posisi ahli hanya untuk memperkuat data empiris dari hasil pengujian.

### 3.6.2 Analisis Faktor

Metode statistik multivariat untuk mengevaluasi hubungan antara variabel yang diamati adalah analisis faktor. Charles Spearman, seorang psikolog yang tertarik untuk memahami kompleksitas kecerdasan manusia, mengembangkan

teknik ini pada awal tahun 1900-an. Spearman menemukan bahwa kemampuan kognitif yang berbeda saling terkait dan dapat diekspresikan secara efisien dalam satu faktor yang mendasarinya, "Fungsi Intelektual bersama yang hebat," juga dikenal sebagai kecerdasan umum atau faktor g (Albani, 2020).

Tujuan utama analisis faktor adalah mendefinisikan struktur suatu data matriks dan menganalisis struktur saling berhubungan (korelasi) antara sejumlah besar variabel (*test score*, *test items*, jawaban kuesioner) dengan cara mendefinisikan satu set kesamaan variabel atau dimensi dan sering disebut dengan faktor atau komponen. Dengan analisis faktor atau komponen, peneliti mengidentifikasi dimensi suatu struktur kemudian menentukan sampai seberapa jauh setiap variabel dapat dijelaskan oleh setiap dimensi. Begitu dimensi dan penjelasan setiap variabel diketahui, maka dua tujuan utama analisis faktor dapat dilakukan yaitu summarization dan data reduction.

#### **a. Asumsi Analisis Faktor**

Analisis faktor menghendaki bahwa matriks data harus memiliki korelasi yang cukup agar dapat dilakukan analisis faktor. Jika berdasarkan data visual tidak ada nilai korelasi yang di atas 0,30, maka analisis faktor tidak dapat dilakukan (Ghozali, 2021). Korelasi antar variabel dapat juga dianalisis dengan menghitung *partial correlation* antar variabel yaitu korelasi antar variabel dengan asumsi variabel lainnya dianggap konstan. Cara lain untuk menentukan dapat atau tidaknya dilakukan analisis faktor adalah dengan melihat matriks korelasi keseluruhan. Untuk menguji apakah terdapat korelasi antar variabel digunakan uji *Bartlett test of sphericity*. Jika hasilnya signifikan berarti matriks korelasi memiliki korelasi signifikan dengan jumlah variabel. Uji lain yang digunakan untuk melihat interkorelasi antar variabel dan dapat tidaknya analisis faktor dilakukan adalah *measure of sampling adequacy* (MSA). Nilai MSA bervariasi dari 0 sampai 1, jika nilai MSA < 0,50, maka analisis faktor tidak dapat dilakukan.

## b. Tahapan Analisis Faktor

Langkah pertama ini dilakukan dengan mencari korelasi matriks antara indikator-indikator yang diobservasi. Metode yang digunakan dalam mencari matriks korelasi antar indikator pada penelitian ini adalah metode Kaiser-Meyer Olkin (KMO). Metode ini banyak digunakan untuk melihat syarat kecukupan data analisis faktor. Metode KMO ini mengukur kecukupan sampling secara menyeluruh dan mengukur kecukupan sampling untuk setiap indikator. Metode ini mengukur homogenitas indikator. Metode KMO ini tidak memerlukan uji statistika, tetapi ada petunjuk yang bisa digunakan untuk melihat homogenitas indikator. Adapun formula yang digunakan untuk mengukur KMO (Widarjono, 2020) adalah:

$$KMO = \frac{\sum \sum r_{ij}^2}{\sum \sum r_{ij}^2 + \sum \sum a_{ij}^2}$$

Dimana:

$r_{ij}^2$  = koefisien korelasi

$a_{ij}^2$  = koefisien korelasi parsial

**Tabel 3. 4 Ukuran KMO**

Ukuran KMO	Rekomendasi
$\geq 0,90$	Sangat baik (Marvelous)
<b>0,80 – 0,89</b>	Berguna (Meritorius)
<b>0,70 – 0,79</b>	Biasa (Middling)
<b>0,60 – 0,69</b>	Cukup (Mediocre)
<b>0,50 – 0,59</b>	Buruk (Miserable)
$\leq 0,50$	Tidak diterima (Unacceptable)

Secara umum tingginya KMO sangat diperlukan. Dari Tabel 3.4 tersebut dapat disarankan untuk paling tidak di atas 0,80. Namun di atas 0,50 biasanya masih bisa diakomodasi untuk penentuan analisis faktor. Selain memasukkan semua indikator di dalam perhitingan korelasi, Kaiser-Meyer Olkin juga menghitung koefisien korelasi di dalam analisis faktor untuk indikator tertentu dengan formula sebagai berikut.

$$KMO = \frac{\sum r_{ij}^2}{\sum r_{ij}^2 + \sum a_{ij}^2}$$

Dimana:

$r_{ij}^2$  = koefisien korelasi

$a_{ij}^2$  = koefisien korelasi parsial

Sebagaimana KMO, semakin tinggi koefisien korelasi MSA maka sangat beralasan untuk memasukka indikator secara individual di dalam analisis faktor.

### 3.6.3 Confirmatory Factor Analisis (CFA)

Analisis faktor konfirmatori, atau CFA, digunakan untuk mengonfirmasi struktur faktor yang sesuai untuk suatu instrumen (Albani, 2020). Analisis faktor konfirmatori digunakan untuk menentukan apakah konstruk yang diajukan oleh suatu teori ada dan dapat dibedakan. Dengan kata lain, konstruksi adalah titik awal untuk analisis faktor konfirmatori. Itu diakhiri dengan mengkonfirmasi atau tidak mengkonfirmasi keberadaan dari sebuah instrumen, sedangkan analisis faktor dimulai dan diakhiri dengan menemukan konstruksi ( Gall & Gall, 2007). Adapun langkah-langkah dasar dari CFA (Albani, 2020) adalah:

1. Membuat hipotesis struktur faktor yang diusulkan dengan kendala yang diperlukan,
2. Mengumpulkan data dan menyesuaikan model CFA,
3. Evaluasi model dan pengujian statistik untuk kecocokan model
4. Merevisi model, membandingkannya dengan model yang kurang lebih kompleks, dan
5. Mmengulangi evaluasi dan pengujian kecocokan sesuai kebutuhan.

### Uji Kelayakan Model

#### 1) *Chi Squares* ( $\chi^2$ )

*Chi squares* digunakan untuk menguji kelayakan model analisis faktor konfirmatori. Hipotesis nol dalam uji chi squares merupakan perbedaan antara sampel dan matriks kovarian yang diestimasi adalah nol sedangkan hipotesis alternatifnya menyatakan ada perbedaan antara sampel dan matriks kovarian yang diestimasi. Nilai df untuk uji chi squares ini besarnya sama dengan jumlah elemen kovarian matriks yang tidak sama dengan jumlah a-parameter yang diestimasi. Jika

nilai chi squares kritis maka kita menolak hipotesis nol dan sebaliknya jika nilai chi squares lebih kecil dari chi squares kritisnya maka kita menerima hipotesis nol. Jika nilai *p-value* lebih kecil  $\alpha$  maka kita menolak hipotesis nol dan sebaliknya jika *p-value* lebih besar dari  $\alpha$  maka gagal menolak hipotesis nol. Jika kita menerima hipotesis nol atau menolak hipotesis alternatif berarti kita dapat menyimpulkan bahwa tidak ada perbedaan antara sample dan matriks kovarian. Artinya model yang kita pilih adalah layak. Sedangkan bila kita menolak hipotesis nol atau menerima hipotesis alternatif maka model tidak layak.

## 2) *Root Mean Squares Error of Approximation (RMSEA)*

Kelemahan uji chi squares adalah sangat sensitif terhadap jumlah sampel. Sebagai alternatif dan pembanding uji chi squares telah dikembangkan uji kelayakan analisis faktor konfirmatori. Salah satunya adalah Root Means Squares Error of Approximation (RMSEA). Adapun formula dari RMSEA sebagai berikut.

$$RMSEA = \sqrt{\frac{\chi^2 - p(p+q)/2 - q}{(n-1)p(p+q)/2 - q}}$$

Dimana:  $\chi^2$  = nilai  $\chi^2$  model; p = jumlah variabel indikator; q = jumlah paramet yang diduga; n jumlah sampel. Sebagai rule of thumb untuk menguji kelayakan model, cutt off value adalah bila  $RMSEA \leq 0,080$ . Jika nilai RMSEA besarnya 0,080 atau lebih kecil maka model dianggap layak. Sebaliknya jika nilainya di atas 0,080 maka model dianggap tidak layak.

## 3) *Goodness of Fit Index (GFI)*

Uji kelayakan model CFA juga dievaluasi dengan menggunakan Goodness of Fit Index (GFI). GFI dihitung dengan menggunakan formula sebagai berikut.

$$GFI = 1 - \frac{tr[(\Sigma^{-1}S - I)^2]}{r[(\Sigma^{-1}S)^2]}$$

Dimana tr = trace matriks; S = kovarian matriks awal;  $\Sigma$  = kovarians matriks model; I = identitas matriks.

Uji kelayakan GFI ini seperti nilai koefisien determinasi ( $R^2$ ) di dalam uji kelayakan atau kebaikan hasil regresi, nilai  $0 \leq GFI \leq 1$ . Semakin tinggi nilai GFI atau

mendekati 1 maka semakin layak model sedangkan nilai GFI semakin mendekati model maka semakin tidak layak model. Sebagai *rule of thumb* biasanya model dianggap layak bila nilai  $GFI \geq 0,90$  sebagai *cut off value*-nya.

#### 4) *Adjusted Goodness of Fit Index (AGFI)*

Uji kelayakan *Adjusted Goodness of Fit Index* (AGFI) merupakan uji kelayakan GFI yang disesuaikan. AGFI ini analog dengan koefisien determinasi yang disesuaikan ( $\text{adjusted } R^2$ ) dalam regresi berganda. AGFI ini merupakan GFI yang disesuaikan dengan derajat kebebasan (*degree of freedom*). Adapun formula untuk AGFI sebagai berikut.

$$AGFI = 1 \left[ \frac{p(p+1)}{sdf} \right] [1 - GFI]$$

Dimana  $p$  = jumlah indikator;  $df$  = *degree of freedom*.

Nilainya terletak antara  $0 \leq GFI \leq 1$ . Sebagaimana uji kelayakan GFI, semakin nilainya mendekati 1 maka semakin baik model dan sebaliknya semakin mendekati angka 0 maka semakin tidak layak model. Namun, tidak ada nilai yang pas AGFI untuk menentukan apakah model layak. Sebagai *rule of thumb, cut off value* adalah bila  $AGFI \geq 0,80$  sebagai model yang layak (*goodness of fit*).

#### 5) *Root Means Squares Residual (RMSR)*

Root Means Squares Residual (RMSR) merupakan akar rata-rata pangkat residual. Adapun formula untuk mencari uji kelayakan model RMSR sebagai berikut.

$$RMSR = \sqrt{\frac{\sum_{i=1}^p \sum_{j=1}^i (S_{ij} - \hat{\sigma}_i)^2}{p(p+1)/2}}$$

Semakin kecil nilai RMSR model semakin sesuai (Fit) atau layak karena ada kesesuaian antara model dan data, sebaliknya semakin besar RMSR model semakin tidak sesuai atau kurang layak. Para peneliti menggunakan *cut off value* sebesar 0,05. Jika nilai RMSR sama atau kurang dari 0,05 maka model adalah baik (fit) sedangkan jika nilainya lebih dari 0,05 maka model kurang fit.

## 6) Uji Signifikansi

Uji t digunakan untuk mengevaluasi signifikansi parameter estimasi, dalam hal ini membandingkan nilai t hitung dengan nilai kritisnya. Jika t hitung lebih besar dari nilai t kritisnya berarti signifikan, sebaliknya bila nilai t hitung lebih kecil dari nilai kritisnya maka tidak signifikan. Besarnya nilai  $\alpha$  biasanya atau secara konvensional ditetapkan sebesar 5% (0,05). Jika nilai t hitung lebih besar dari  $\pm 1,96$  maka variabel indikator dikatakan signifikan dan jika tidak maka tidak signifikan.

## 7) *Squared Multiple Correlation (R<sup>2</sup>)*

Square multiple Correlation dilakukan setelah uji signifikansi parameter untuk melihat seberapa besar varian variabel laten menjelaskan variabel indikator. Total varian dari setiap variabel laten dapat dipilah menjadi dua yaitu varian yang berhubungan dengan variabel laten dan varian yang berhubungan faktor spesifik yang berasal dari error atau residual. Proporsi dari varian yang berhubungan dengan faktor laten ini disebut dengan *communality*. Dengan demikian semakin besar square multiple correlation semakin bisa dipercaya (*more reliable*) variabel indikator sebagai pengukur variabel laten. Sebaliknya semakin kecil square multiple correlation semakin tidak bisa dipercaya (*less reliable*) variabel indikator sebagai pengukur variabel laten.

## 8) Respesifikasi Model

Langkah terakhir dari CFA adalah respesifikasi model. Respesifikasi model ini dilakukan dengan menggunakan ukuran-ukuran diagnosis dan teori-teori yang melandasi adanya respesifikasi model. Kemungkinan yang bisa dilakukan agar bisa memenuhi kelayakan model adalah dengan cara:

- a. Melakukan korelasi antara variabel indikator.
- b. Melakukan korelasi antara variabel residual.

### 3.6.4 Pengolahan Data dan Uji Hipotesis Hasil Implementasi Instrument Test *Creative Thinking*

Data kelas eksperimen diperoleh dari *pretest* dan *posttest*. Setelah diperoleh data dari kedua kelas tersebut maka dilakukan langkah-langkah sebagai berikut:

#### 1. Penskoran

Penskoran tes pilihan ganda dan essay dilakukan dengan menggunakan pedoman penskoran. Sebelum lembar jawaban siswa diberi skor, terlebih dahulu ditentukan standar penilaian untuk tiap tahap sehingga dalam pelaksanaannya unsur subjektifitas dapat dikurangi. Skor setiap siswa ditentukan dengan menghitung jumlah jawaban yang benar. Pemberian skor dihitung dengan menggunakan rumus:

$$S = \Sigma R$$

Dengan: S= Skor siswa dan R= jawaban siswa yang benar

#### 2. Mengubah skor mentah menjadi nilai dilakukan dengan mengacu pada Penilaian Acuan Patokan (PAP).

$$\text{Nilai} = \frac{\text{skor mentah}}{\text{skor maksimum ideal}} \times 100$$

#### 3. Menghitung nilai maksimum, minimum, dan rata-rata hasil pre-test dan post-test.

#### 4. Menghitung nilai N-Gain dengan menggunakan rumus Hake (Kusnendi, 2013) sebagai berikut:

$$\text{Gain ternormalisasi (g)} = \frac{\text{skor posttest} - \text{skor pretest}}{\text{skor maksimal} - \text{skor pretest}}$$

Keterangan:

(g) = gain yang dinormalisir

*Posttest* = tes diakhir pembelajaran

*Pretest* = tes diawal pembelajaran

Acuan kriteria perolehan gain yang sudah dinormalisasikan sebagai berikut:

**Tabel 3. 5 Kriteria Indeks Gain**

<b>Skor</b>	<b>Katagori</b>
$(g) \geq 0,70$	<b>Tinggi</b>
$0,30 \leq (g) < 0,70$	<b>Sedang</b>
$(g) < 0,30$	<b>Rendah</b>

### 5. Uji Hipotesis

Uji hipotesis penelitian didasarkan pada data kemampuan berpikir kreatif yaitu data selisish nilai *pre test* dan *post test*. Pengujian hipotesis dilakukan dengan menggunakan uji t-independen dua arah (*t-test independent*). Uji t independen dua arah ini digunakan untuk menguji signifikasi perbedaan rata-rata (*mean*) dua kelompok sampel eksperimen yang tidak berhubungan. Rumus yang digunakan adalah :

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

Keterangan :

$\hat{Y}_1$  dan  $\hat{Y}_2$  = nilai rata-rata sampel

$S_1^2$  dan  $S_2^2$  = varians sampel

$n_1$  dan  $n_2$  = ukuran sampel