

### **BAB III**

#### **METODE PENELITIAN**

##### **A. Jenis Penelitian**

Penelitian ini merupakan jenis penelitian *quasy experiment* dengan desain non-equivalent control group dan *posttest only control group*. Pada desain ini peneliti menggunakan dua kelas yang berbeda dengan perlakuan yang berbeda pula, yang terdiri dari kelas eksperimen (E) dan kelas kontrol (K). Kelas eksperimen (E) merupakan rombongan belajar yang diberikan perlakuan berupa pendekatan pembelajar saintifik (5M) berbasis analogi, sedangkan kelas kontrol (K) merupakan rombongan belajar yang menerapkan pendekatan pembelajar saintifik (5M). Masing-masing kelas akan diberikan tes pada akhir proses pembelajaran (*posttest*) yakni setelah perlakuan selesai diberikan. Bentuk tes soal yang diberikan berupa soal uraian keterampilan berpikir tingkat tinggi yang mencakup materi kesetimbangan kimia. Data hasil *posttest* akan dianalisis lebih lanjut mengenai keterampilan berpikir tingkat tinggi siswa. Desain penelitian ini disajikan seperti pada Tabel 2

**Tabel 2. Desain *Posttest only Control Group***

<b>Kelompok</b>	<b>Perlakuan</b>	<b><i>Posttest</i></b>
Eksperimen	X	Q
Kontrol	Y	Q

Keterangan:

X : Pembelajaran kimia dengan pendekatan saintifik (5M) berbasis analogi

Y : Pembelajaran kimia dengan pendekatan saintifik (5M)

Q : Tes Keterampilan berpikir tingkat tinggi (HOTS)

## **B. Waktu dan Tempat Penelitian**

Penelitian ini dilaksanakan pada bulan November 2018 di SMA N 1 Seyegan Yogyakarta dengan pembahasan materi kesetimbangan kimia

## **C. Populasi dan Sampel Penelitian**

### **1. Populasi**

Populasi penelitian yang digunakan adalah siswa kelas XI SMA di Daerah Istimewa Yogyakarta yang setara dengan SMA Negeri 1 Seyegan Yogyakarta. Ciri-ciri SMA Negeri 1 Seyegan adalah sebagai berikut

- a. Menerapkan kurikulum 2013
- b. Terakreditasi A
- c. Guru mata pelajaran kimia yang berkompeten di bidang pendidikan kimia
- d. Jumlah siswa satu rombongan belajar paling sedikit 20 siswa dan paling banyak 36 siswa
- e. Memiliki sarana dan prasarana yang memadai, seperti laboratorium kimia, fisika, biologi, komputer dan perpustakaan

### **2. Sampel**

Sampel yang digunakan pada penelitian ini adalah dua kelas siswa kelas XI SMA Negeri 1 Seyegan dengan teknik pengambilan sampel *simple random sampling*

## **D. Variabel Penelitian**

Penelitian ini memiliki dua variabel, yaitu variabel bebas dan variabel terikat. Variabel bebas pada penelitian ini adalah pendekatan saintifik (5M) berbasis

analogi dan pendekatan saintifik (5M), sedangkan variabel terikatnya adalah keterampilan berpikir tingkat tinggi. Penjelasan definisi operasional masing-masing variabel penelitian dapat dilihat pada Tabel 3 berikut ini

**Tabel 3. Definisi Operasioanl Variabel Penelitian**

<b>Variabel Penelitian</b>	<b>Definisi Operasional</b>
Pendekatan saintifik	Pendekatan saintifik adalah pendekatan ilmiah yang digunakan sebagai konsep pembelajaran di dalam kurikulum 2013 yang dijabarkan ke dalam lima praktek kegiatan pembelajaran yaitu, Mengamati, Menanya, Mencari Informasi, Mengasosiasi, dan Mengkomunikasikan
Pendekatan saintik berbasis analogi	Pendekatan saintifik berbasis analogi adalah pendekatan ilmiah yang memiliki unsur analogi dalam tahapan pembelajarannya.
Keterampilan berpikir tingkat tinggi	Keterampilan berpikir tingkat tinggi adalah keterampilan berpikir siswa untuk menghubungkan berbagai informasi yang diperoleh menjadi serangkaian informasi yang utuh. Kemampuan ini diukur dengan melihat kemampuan siswa pada tiga ranah kognitif. Ranah kognitif yang dimaksud adalah C4 (Menganalisis), C5 (Mengevaluasi), dan C6 (Mencipta)

#### **E. Penerapan Pembelajaran Kimia berbasis Analogi**

Implementasi pembelajaran kimia berbasis analogi pada penelitian ini dikembangkan berdasarkan aturan FAR (*Focus* (Fokus), *Action* (Aksi), *Reflection* (Refleksi) sebagaimana tercantum pada Tabel 4 sebagai berikut

**Tabel 4. Langkah-langkah Analogi dengan Aturan FAR**

<b>Fokus</b>		
Konsep	Menjelaskan konsep kesetimbangan dinamis secara umum	
Siswa	Siswa menjelaskan apa yang diketahui tentang kesetimbangan dinamis	
Analogi	Guru menjelaskan analog yang digunakan	
<b>Aksi</b>		
Mirip	Menjelaskan bagian-bagian analog yang mirip dengan konsep target	Menjelaskan konsep target yang ingin dicapai
Tidak Mirip	Menjelaskan bagian-bagian analogi yang tidak mirip dengan konsep target	
<b>Refleksi</b>		
Kesimpulan	Apakah analogi jelas dan berguna?	
Perbaikan	Apakah ada perubahan yang perlu anda lakukan pada saat anda menggunakan analogi ini?	

## **F. Teknik dan Instrumen Pengumpulan Data**

### **1. Teknik Pengumpulan Data**

Teknik pengumpulam data yang digunakan adalah teknik tes dan non tes. Bentuk tes yang digunakan adalah ujian tulis dengan instrumen penilaian berupa uraian bebas dengan materi kesetimbangan kimia. Ujian uraian bebas diberikan pada saat *posttest*, sedangkan bentuk non tes yang digunakan adalah observasi dengan instrumen penilaian berupa lembar observasi.

### **2. Instrumen Pengumpulan Data**

#### **a. Instrumen Soal Keterampilan Berpikir Tingkat Tinggi**

Instrument soal keterampilan berpikir tingkat tinggi adalah instrument yang digunakan untuk pengumpulan data dalam penelitian ini.

Kisi-kisi instrumen tes keterampilan berpikir tingkat tinggi menurut

Anderson dan Krathwohl (2001) secara lengkap terdapat pada Lampiran 5 dan secara ringkas tercantum pada Tabel 5.

**Tabel 5. Kisi-kisi Instrumen Tes Keterampilan Berpikir Tingkat Tinggi**

Materi	Indikator Materi	Indikator HOTS									Jumlah Soal
		C4			C5		C6				
		Sub indikator			Sub Indikator		Sub Indikator				
		a*	b*	c*	d*	e*	f*	g*	h*		
Keseimbangan Dinamis	Menjelaskan keseimbangan dinamis	1				3				3	
		2									
Hukum Tetap dan Tetap Kesetimbangan	Mengolah data untuk menentukan nilai tetap kesetimbangan ( $K_C$ dan $K_p$ )	6			4					7	
		7			5						
		8			9	10					
Kesetimbangan	Menganalisis data untuk menentukan hubungan nilai tetap kesetimbangan	12	11							2	
		13									2
	Menganalisis data dalam perhitungan system kesetimbangan disosiasi	14									
Faktor-Faktor yang memengaruhi Kesetimbangan Kimia	Menentukan pergeseran kesetimbangan kimia berdasarkan perubahan konsentrasi, tekanan, volume, dan suhu	19			15	20	16	21		8	
					17						
					18						
					22						

Keterangan: a). Membedakan, b). Mengorganisasikan, c). Mengatribusikan, d). Memeriksa, e). Mengkritik, f). Merumuskan, g). Merencanakan, h). Memproduksi

Instrumen keterampilan berpikir tingkat tinggi pada penelitian ini berlandaskan teori ranah kognitif yang dikemukakan oleh Anderson dan Krathwohl (2001) khususnya pada ranah kognitif menganalisis, mengevaluasi, dan mencipta.

### **b. Instrumen Lembar Kerja Siswa (LKS)**

Lembar Kerja Siswa (LKS) merupakan instrumen yang disusun untuk memandu siswa selama proses pembelajaran disertai dengan menjawab pertanyaan terkait dengan sub materi yang dipelajari. LKS disusun sesuai dengan pendekatan pembelajaran yang diterapkan yaitu pendekatan pembelajaran saintifik berbasis analogi untuk kelas eksperimen dan pendekatan saintifik untuk kelas kontrol.

### **c. Lembar Pengamatan Keterlaksanaan Pembelajaran Kimia berbasis Analogi**

Aspek pengamatan keterlaksanaan pembelajaran kimia berbasis analogi secara lengkap tercantum pada Lampiran 7, dan secara garis besar dapat dilihat pada Tabel 6. Lembar observasi dan tanya jawab keterlaksanaan pembelajaran digunakan untuk mengetahui persentase keterlaksanaan penerapan analogi dalam pembelajaran. Lembar penilaian ini diisi oleh pengamat (*observer*) selama proses pembelajaran berlangsung. Pengamat akan memberikan tanda ceklis (√) pada kolom Ya jika suatu aspek yang dinilai berhasil terlaksana dan memberikan tanda silang (X) jika suatu aspek yang dinilai tidak berhasil terlaksana. Lembar penilaian ini disusun berdasarkan teori yang telah ada dan dikembangkan oleh peneliti sehingga setiap aspek terbagi menjadi beberapa sub aspek penting.

**Tabel 6. Aspek Pengamatan Keterlaksanaan Pembelajaran Kimia berbasis Analogi**

No	Aturan FAR	Aspek	Sub Aspek	Teknik Pengumpulan Data
1	Fokus	Konsep	1. Kebenaran konsep	Observasi
			2. Sistematis	Observasi
			3. Jelas	Observasi
			4. Mencakup kedalaman dan keluasan konsep	Observasi
			5. Sesuai dengan indikator pencapaian kompetensi	Observasi
		Siswa	1. Mengikuti pelajaran dengan baik	Observasi
			2. Berdiskusi dengan teman	Observasi
			3. Memahami konsep materi	Tanya Jawab
			4. Antusias	Observasi
		Analogi	1. Analog familiar bagi siswa	Tanya Jawab
2. Hubungan analog dengan target konsep dapat diterima oleh siswa	Tanya Jawab			
2	Aksi	Mirip	Menjelaskan analog yang mirip dengan target	Observasi
		Tidak Mirip	Menjelaskan analog yang tidak mirip dengan target	Observasi
3	Refleksi	Kesimpulan	1. Analog memudahkan siswa memahami konsep	Tanya Jawab
			2. Analog jelas dan berfungsi dengan baik	Tanya Jawab
		Perbaikan	Analogi tidak perlu diperbaiki	Tanya Jawab

Teknik dan instrumen pengumpulan data yang telah disebutkan didukung oleh perangkat pembelajaran yang berupa Rencana Pelaksanaan Pembelajaran (RPP) yang disusun berdasarkan silabus yang ada di sekolah SMA Negeri 1

Seyegan Yogyakarta. Perangkat pembelajaran dan seluruh instrumen pengumpulan data yang telah disusun, sebelum diterapkan dalam penelitian, seluruhnya divalidasi teoritis oleh dua dosen ahli dan terdapat penambahan enam validator lainnya untuk instrument soal HOTS. Hasil validasi teoritis selanjutnya direvisi dan untuk instrumen soal HOTS dilanjutkan dengan validasi empiris, yaitu dengan melakukan uji coba kepada siswa-siswa di luar sampel penelitian yang telah mempelajari materi kesetimbangan kimia. Hasil uji coba selanjutnya dianalisis untuk menghasilkan soal HOTS yang layak digunakan.

#### **G. Pemodelan *Rasch***

Fitur utama dari model Rasch adalah tabel probabilitas respons terhadap item. Probabilitas keberhasilan tergantung pada perbedaan antara kemampuan orang dan kesulitan item (Bond dan Fox, 2015: 10). Menurut Lord (1980) di dalam Nurcahyo (2016) model ini mengasumsikan bahwa semua aitem mendiskriminasikan secara sama serta tidak dapat dijawab dengan benar berdasarkan tebakan. Bond dan Fox (2015: 32-37) menyatakan bahwa ada beberapa asumsi yang harus diperhatikan, yakni unidimensi, *item fit*, dan abilitas (*person* dan *item*). Pemodelan Rasch merupakan sebuah model analisis yang digunakan pada penelitian ini, yang berfungsi untuk menganalisis kesukaran butir soal pada setiap ranah kognitifnya. Pemodelan Rasch akan dilakukan jika memenuhi beberapa syarat sebagai berikut

## 1. Asumsi Unidimensi

Dalam teori umum sifat laten, mengasumsikan bahwa seperangkat sifat laten mendasari kinerja peserta ujian pada satu set item tes. Hanya ada satu kemampuan untuk menjelaskan atau memperhitungkan kinerja peserta ujian. Respon butir yang mengasumsikan kemampuan laten tunggal atau setiap butir hanya mengukur satu kemampuan disebut unidimensional. Uji unidimensi dilakukan dengan memperhatikan nilai *raw variances explained by measures* hasil *output* pada tabel *item unidimensionality* dengan kriteria seperti yang ada pada Tabel 7 (Fisher, 2007)

**Tabel 7. Kriteria Hasil Uji Unidimensi**

<b>Nilai <i>raw variances explained by measures</i></b>	<b>Kriteria</b>
<50%	Lemah
50-60%	Jelek
60-70%	Bagus
70-80%	Sangat Bagus
>80%	Sempurna

Hal lain yang mendukung asumsi unidimensi adalah dengan memperhatikan nilai keragaman yang tidak dapat dijelaskan (*unexplained variance*) dengan kriteria penilaian seperti pada Tabel 8 berikut

**Tabel 8. Kriteria *unexplained variance***

<b>Nilai <i>unexplained variance</i></b>	<b>Kriteria</b>
>15%	Lemah
10-15%	Jelek
5-10%	Bagus
3-5%	Sangat Bagus
<3%	Sempurna

Data hasil analisis dapat dilihat pada Tabel 9. Nilai persentase *empirical* yang diperoleh adalah 60,1%. Berdasarkan kriteria asumsi unidimensi maka kriteria yang diperoleh adalah bagus, sehingga asumsi unidimensionalitas terpenuhi. Data ini juga didukung dengan adanya data *unexplained variance* yang berada diantara 2,7%-8,1% yang menunjukkan tingkat independensi item dalam instrument yang baik. Hasil analisis secara lengkap dapat dilihat pada Lampiran 9

**Tabel 9. Standardized Residual Variance (in Eigenvalue units)**

<i>Variance</i>	<i>Empirical</i>			<i>Modeled</i>
<i>Total raw variance in observations</i>	47,7	100%	-	100%
<i>Raw variance explained by measures</i>	28,7	60,1%	-	59,8%
<i>Raw variance explained by persons</i>	4,1	8,7%	-	8,6%
<i>Raw variance explained by items</i>	24,5	51,5%	-	51,2%
<i>Raw unexplained variance (total)</i>	19	39,9%	100%	40,2%
<i>Unexplained variance in 1<sup>st</sup> contrast</i>	3,9	8,1%	20,3%	-
<i>Unexplained variance in 2<sup>nd</sup> contrast</i>	2,1	4,5%	11,2%	-
<i>Unexplained variance in 3<sup>rd</sup> contrast</i>	1,9	4,1%	10,2%	-
<i>Unexplained variance in 4<sup>th</sup> contrast</i>	1,5	3,2%	8,1%	-
<i>Unexplained variance in 5<sup>th</sup> contrast</i>	1,3	2,7%	6,8%	-

## 2. Item Fit

Dalam pengukuran Rasch, konsep *fit* adalah sebagai mekanisme kualitas yang mengindikasikan bahwa kualitas instrument yang disusun telah memadai. Hal ini juga digunakan untuk menilai makna konstruk unidimensional, artinya

indeks *fit* membantu peneliti memastikan bahwa persyaratan Rasch untuk unidimensi berlaku secara empiris

Menurut Boone, Staver, dan Yale (2014: 166-173) nilai kriteria yang digunakan untuk memeriksa kesesuaian butir adalah sebagai berikut

- 1) Nilai Outfit MNSQ yang diterima:  $0,5 < \text{MNSQ} < 1,5$
- 2) Nilai Outfit ZSTD yang diterima:  $-2,0 < \text{ZSTD} < +2,0$
- 3) Nilai Point Measure Correlation (Pt Mean Corr):  $0,4 < \text{Pt Measure Corr} < 0,85$

**Tabel 10. Item Fit**

No. Soal	Nilai <i>Outfit</i>		
	<i>MNSQ</i>	<i>ZSTD</i>	<i>Pt-Measure Corr</i>
2	3,09	9,9	0,44
18	1,37	1,8	0,31
16	1	0,0	0,34
14	0,91	-0,1	0,16
7	0,75	-0,5	0,18
10	0,97	-0,1	0,25
12	0,79	-0,9	0,31
21	0,88	-1	0,33
4	0,87	-1,1	0,22
19	0,87	-1,1	0,40
11	0,84	-1,5	0,49
13	0,72	-1,3	0,39
20	0,80	-1,8	0,49
5	0,79	-1,6	0,40
1	0,76	-2	0,49
3	0,76	-2	0,49
17	0,73	-2,4	0,38
22	0,71	-2,5	0,34
9	0,34	-7,3	0,56

Pada program *winstep* untuk keperluan analisis ini dapat dilihat pada tabel *output Item Fit Order* seperti yang ada pada Tabel 10. Berdasarkan hasil analisis soal dengan nomor 6, 8, dan 15 tidak terdeteksi di dalam hasil analisis *output Item*

*Fit Order*, hal ini menunjukkan bahwa soal tersebut merupakan soal dengan kategori yang sangat sulit, sehingga otomatis soal tersebut menjadi soal yang tidak *fit*. Hasil analisis secara lengkap dapat dilihat pada Lampiran 9

Berdasarkan hasil analisis yang tertera pada Tabel 13, soal-soal yang tidak *fit* adalah butir soal dengan nomor 2, 9, 17, dan 22, hal ini disebabkan karena karena pada soal tersebut terdapat dua kriteria yang tidak terpenuhi. Soal nomor 2 dan 9 tidak memenuhi kriteria *outfit MNSQ* dan *outfit ZSTD*, sedangkan soal nomor 17 dan 22 tidak memenuhi kriteria *outfit ZSTD* dan *pt-measure corr*. Hasil analisis tidak menunjukkan untuk soal nomor 6, 8, dan 15, hal ini disebabkan karena pada hasil analisis tingkat kesukaran, soal tersebut berada pada posisi tingkat paling sukar, sehingga tidak ada yang mampu berhasil menyelesaikan soal tersebut.

Butir soal nomor 4, 7, 10, 12, 13, 14, 16, 18, 19, dan 21 dalam kriteria *pt-measure corr* terlihat tidak terpenuhi, namun soal-soal tersebut masih dapat digunakan karena masih memenuhi dua kriteria lainnya. Menurut Sumintono dan Widhiarso (2015: 72) soal dikategorikan *fit* dengan model minimal dua kriteria harus terpenuhi, sehingga dari 22 soal dengan 144 responden, terdapat 7 soal yang tidak valid dan 15 soal valid. Soal yang tidak valid tersebut tidak layak digunakan untuk tujuan analisis penelitian ini. Ringkasan nomor butir soal yang valid dan tidak valid dapat dilihat pada Tabel 11

**Tabel 11. Nomor Soal Valid dan Tidak Valid**

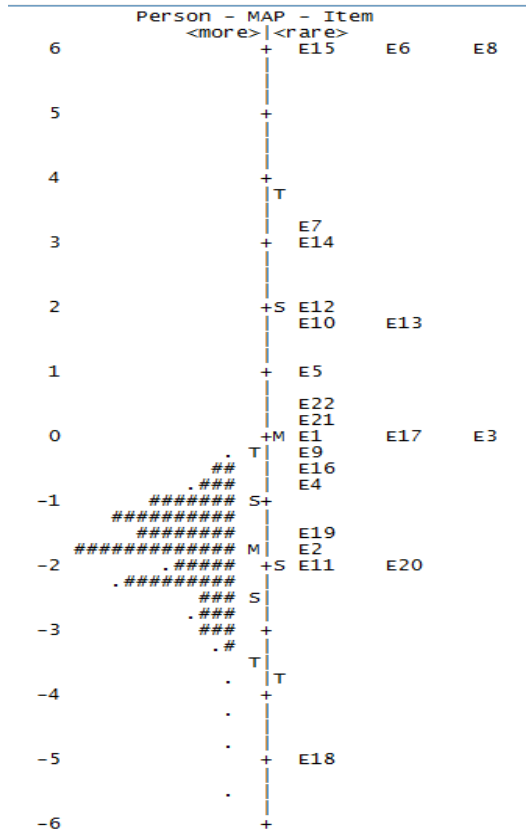
<b>Nomor Soal Valid</b>	<b>Nomor Soal Tidak Valid</b>
1, 3, 4, 5, 7, 10, 11, 12, 13, 14, 16, 18, 19, 20, dan 21	2, 6, 8, 9, 15, 17, dan 22

### 3. Abilitas/Estimasi Pengukuran

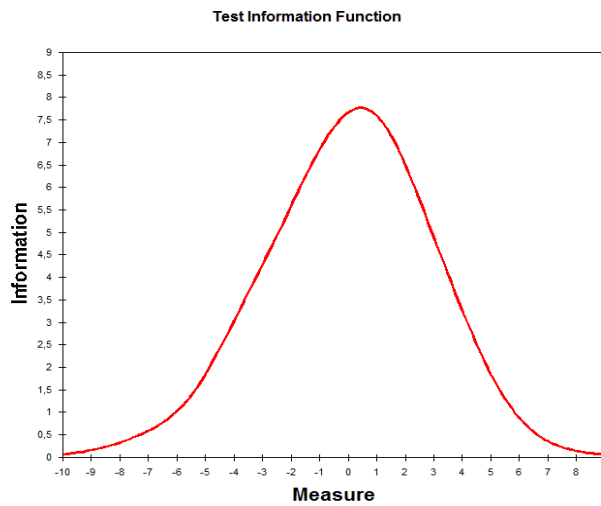
Dalam konteks penilaian pendidikan, penting diketahui abilitas kemampuan siswa dan abilitas butir soal. Grafik distribusi tingkat kesukaran butir dan kemampuan responden secara bersamaan terdapat pada skala logit. Sebelah kiri yang berupa tanda (#) mewakili responden dan sebelah kanan berupa (E) menunjukkan distribusi tingkat kesukaran butir. Siswa yang berada pada logit positif (tinggi) merupakan siswa yang berkemampuan tinggi dan item soal tersebut merupakan soal yang sulit, sedangkan siswa yang berada pada logit yang rendah merupakan siswa dengan kemampuan rendah dan item soal tersebut merupakan soal dengan tingkat kesukaran rendah. Aitem dan siswa yang fit berada di antara -2 hingga +2 dengan ukuran sampel 30 hingga 300 (Bond & Fox, 2007: 38-43)

Berdasarkan hasil analisis seperti pada Gambar 2 sebagian besar kemampuan siswa dan soal tes berada di antara skala -2 s/d +2, sehingga keduanya berada pada posisi sedang. Hal ini juga didukung dari hasil fungsi informasi pengukuran seperti pada Gambar 3. Pada level abilitas rendah, informasi yang diperoleh dari pengukuran juga rendah. Pada level abilitas sedang, informasi yang diperoleh dari pengukuran adalah sangat tinggi. Pada level abilitas tinggi, informasi yang diperoleh dari pengukuran adalah rendah. Hal ini

menunjukkan bahwa butir soal kesetimbangan kimia tersebut menghasilkan informasi yang optimal ketika diberikan pada individu yang akan dianalisis keterampilan berpikir tingkat tinggi nya.



Gambar 2. Variable Maps



Gambar 3. Grafik Fungsi Informasi Pengukuran

## H. Validitas dan Reliabilitas Instrumen

### 1. Validitas

Uji validitas instrument dalam penelitian ini dilakukan dengan menggunakan uji validitas teoritis dan empiris. Jenis validitas beserta instrumen yang digunakan dapat dilihat pada Tabel 12

**Tabel 12. Instrumen dan Jenis Validitas Instrumen**

<b>Instrumen</b>	<b>Jenis Validitas</b>
Tes Kemampuan berpikir Tingkat Tinggi	Teoritis dan Empiris
Analogi	Teoritis
Lembar Kerja Siswa	Teoritis
RPP	Teoritis
Lembar Observasi Pelaksanaan Pembelajaran	Teoritis

#### a. Validitas Teoritis

Validitas teoritis dilakukan oleh dua dosen ahli. Instrumen yang telah dikonstruksi tentang aspek-aspek yang akan diukur yang berlandaskan suatu teori, selanjutnya dikonsultasikan dengan para ahli. Para ahli diminta pendapatnya tentang instrumen yang telah disusun berdasarkan pertimbangan dari segi materi, konstruksi, dan bahasa. Tujuan validitas teoritis adalah untuk dilakukan perbaikan jika memang saat penilaian oleh ahli dinilai kurang layak. Hasil validasi teoritis dari dua dosen ahli menyatakan bahwa instrument layak untuk digunakan dengan perbaikan

Validitas teoritis tes keterampilan berpikir tingkat tinggi akan dianalisis lebih lanjut dengan menggunakan indek Aiken'V, sehingga selain dilakukan oleh dua dosen ahli, juga dilakukan oleh enam validator lainnya yang terdiri

dari lima orang teman sejawat dan satu orang praktisi. Para validator diminta untuk menilai validitas butir. Penilaian dilakukan dengan format sebagai berikut

- 1 = Tidak Valid
- 2 = Kurang Valid
- 3 = Cukup
- 4 = Valid
- 5 = Sangat Valid

Setelah masing-masing validator melakukan penilaian, hasil penilaian digabung dalam satu tabel dan dihitung indeks Aiken'V kesepakatan ahli untuk masing-masing butir. Formula Aiken's yang digunakan adalah sebagai berikut (Aiken, 1985).

$$V = \frac{\sum s}{[n(c - 1)]}$$

Keterangan:

V = Validitas

s = r-Io

Io = Angka penilaian validitas yang terendah

c = Angka penilaian validitas tertinggi

r = Angka yang diberikan oleh penilai

Prosedur untuk menentukan koefisien V dimulai dengan peringkat dari satu item oleh n rater atau peringkat m oleh penilai tunggal. Peringkat validitas dapat dibuat pada setiap skala. Kategori validitas tertinggi sebagai c, validitas rendah sebagai lo, dan rating validitas penilai item sebagai r. Nilai r ditransformasikan menjadi  $s = r - lo$  jika  $lo < hi$ , dan ditransformasikan menjadi  $s = lo - r$  jika  $hi < lo$ . Nilai-nilai s untuk semua penilai kemudian ditambahkan di seluruh n rater atau item m untuk menghasilkan S. Ketika

peringkat dari satu item dibuat oleh  $n$  rater, koefisien  $V$  untuk item tersebut dihitung sebagai  $V = S/[n(c-1)]$ . Ketika peringkat item  $m$  dibuat oleh satu penilai, koefisien  $V$  untuk penilai tersebut dihitung sebagai  $V = S/[n(c-1)]$ . Kisaran koefisien  $V$  adalah 0 hingga 1. Nilai yang tinggi menunjukkan bahwa item memiliki validitas konten yang tinggi (ketika peringkat item tunggal dibuat oleh  $n$  rater) atau bahwa satu set item memiliki validitas konten yang tinggi dalam penilaian satu penilai (ketika peringkat dari  $n$  item dibuat oleh satu penilai). Hasil analisis akan diperoleh dengan membandingkan nilai  $V$  yang diperoleh dengan nilai koefisien reliabilitas. Soal dinyatakan valid jika nilai  $V$  yang diperoleh  $>$  dari nilai koefisien validitas (Aiken, 1985).

Berdasarkan tabel koefisien validitas dengan nilai signifikansi 5%, maka setiap butir soal akan dinyatakan valid jika nilai validitas butir soal lebih besar dari 0.75. Hasil analisis dengan menggunakan rumus Aiken'V dari 22 butir soal uraian menunjukkan semua butir memiliki nilai koefisien validitas  $>0.75$ , sehingga dinyatakan valid. Analisis secara lengkap dapat dilihat pada Lampiran 8. Berdasarkan uji validitas isi, perbaikan instrument terdapat pada stuktur kalimat, penulisan notasi ilmiah, penyempurnaan analogi, dan cara perhitungan jawaban soal tes. Setelah proses perbaikan dan penilaian, instrument tersebut dinyatakan layak untuk diterapkan pada penelitian.

#### **b. Validitas Empiris**

Validitas empiris dilakukan dengan mengujikan tes terhadap siswa yang telah mempelajari materi kesetimbangan kimia dan bukan merupakan sampel dalam penelitian ini, yakni kepada siswa kelas XII yang berjumlah 144

siswa. Uji validitas empiris dianalisis dengan menggunakan model *Rasch* dengan berbantuan program *Winstep*.

Analisis validasi empiris serupa dengan analisis *item fit* pada uji asumsi Rasch, sehingga hasil analisis *item fit* adalah hasil yang sama untuk keperluan analisis validasi empiris.

## 2. Reliabilitas

Model Rasch menyediakan indeks yang membantu peneliti untuk menentukan apakah item cukup tersebar di sepanjang kontinum. Indeks reliabilitas item akan menunjukkan penempatan item di sepanjang jalur, jika item yang sama diberikan kepada sampel yang berbeda akan menghasilkan hasil yang sama (Bond & Fox, 2007: 40-41)

Uji reliabilitas dalam penelitian ini dianalisis dengan menggunakan model Rasch dengan memperhatikan nilai *alpha Cronbach*, *person reliability*, dan *item reliability*. Kriteria untuk menginterpretasikan hasil uji reliabilitas tersebut dapat dilihat pada Tabel 13 dan Tabel 14 (Sumintono, 2015: 85).

**Tabel 13. Kriteria Nilai *alpha Cronbach***

Hasil Reliabilitas	Kriteria
< 0,5	Buruk
0,5 – 0,6	Jelek
0,6 – 0,7	Cukup
0,7 – 0,8	Bagus
>0,8	Sangat Bagus

**Tabel 14. Kriteria Reliabilitas *Person* dan *Item***

Hasil Reliabilitas	Kriteria Reliabilitas
<0,67	Lemah
0,67 – 0,80	Cukup
0,81-0,90	Bagus
0,91 – 0,94	Sangat Bagus
>0,94	Istimewa

Berdasarkan hasil uji reliabilitas menunjukkan *person reliability* 0,64 dan *item reliability* 0,99, dengan kesimpulan bahwa konsistensi jawaban dari siswa lemah, namun kualitas butir-butir soal dalam instrumen tersebut reliabilitasnya istimewa. Interaksi antara person dan butir-butir soal secara keseluruhan dilihat dari nilai *alpha Cronbach* dengan nilai 0,81 yang memiliki kriteria bagus. Hasil analisis dapat dilihat pada Gambar 4

SUMMARY OF 144 MEASURED Person								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	34.4	19.4	-1.81	.43	1.02	.0	.96	-.1
S.D.	5.5	1.8	.81	.07	.40	1.0	.37	.9
MAX.	47.0	22.0	-.20	.91	2.85	2.6	2.25	2.2
MIN.	18.0	13.0	-5.61	.36	.30	-2.7	.25	-2.5
REAL RMSE	.48	TRUE SD	.65	SEPARATION	1.33	Person RELIABILITY	.64	
MODEL RMSE	.44	TRUE SD	.68	SEPARATION	1.54	Person RELIABILITY	.70	
S.E. OF Person MEAN = .07								
VALID RESPONSES: 92.6% (APPROXIMATE)								
Person RAW SCORE-TO-MEASURE CORRELATION = .92 (approximate due to missing data)								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .81 (approximate due to missing data)								
SUMMARY OF 19 MEASURED Item								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	247.4	133.3	.00	.17	.95	-1.0	.95	-.8
S.D.	99.2	9.9	1.88	.06	.55	3.2	.54	3.0
MAX.	532.0	144.0	3.18	.35	3.14	9.9	3.09	9.9
MIN.	130.0	116.0	-4.93	.12	.32	-7.9	.34	-7.3
REAL RMSE	.19	TRUE SD	1.87	SEPARATION	9.81	Item RELIABILITY	.99	
MODEL RMSE	.18	TRUE SD	1.87	SEPARATION	10.14	Item RELIABILITY	.99	
S.E. OF Item MEAN = .44								

Gambar 4. Hasil Keluaran Nilai Reliabilitas

Berdasarkan uji asumsi Rasch, kriteria validitas, dan reliabilitas, maka soal yang digunakan sebagai instrument pengukuran keterampilan berpikir tingkat tinggi pada penelitian ini adalah terdiri dari 8 soal yang fit dengan model, mewakili setiap indikator soal, dan mewakili ranah kognitif HOTS. Adapun 8 soal tersebut dapat dilihat pada Tabel 15 berikut

**Tabel 15. Nomor Soal Valid dan Soal untuk Instrumen Penelitian**

Nomor Soal Valid	Nomor Soal untuk Instrumen Penelitian
1, 3, 4, 5, 7, 10, 11, 12, 13, 14, 16, 19, 20, dan 21	1, 5, 10, 12, 13, 16, 19, dan 21

## I. Teknik Analisis Data

Analisis data pada penelitian ini menggunakan uji *independent sample t-test* yang berfungsi untuk mengetahui perbedaan keterampilan berpikir tingkat tinggi siswa antara siswa yang menerapkan sistem pembelajaran kimia berbasis analogi dengan siswa yang menerapkan sistem pembelajaran kimia tanpa analogi.

### 1. Analisis Data

Analisis data pada penelitian dilakukan dengan menggunakan uji t yang bertujuan untuk mengetahui ada atau tidak perbedaan keterampilan berpikir tingkat tinggi siswa pada materi kesetimbangan kimia antara dua sampel yang memperoleh perlakuan yang berbeda. Analisis dilakukan dengan berbantuan program SPSS

Hipotesis statistiknya adalah sebagai berikut.

$H_0$  : Tidak ada perbedaan keterampilan berpikir tingkat tinggi antara siswa yang menerapkan pembelajaran kimia berbasis analogi dengan siswa yang tidak menerapkan pembelajaran kimia berbasis analogi

$H_a$ : Ada perbedaan keterampilan berpikir tingkat tinggi antara siswa yang menerapkan pembelajaran kimia berbasis analogi dengan siswa yang tidak menerapkan pembelajaran kimia berbasis analogi

Kriteria pengujiannya adalah jika nilai signifikansi  $\leq 0,05$  maka hipotesis nol ditolak (Pallant, 2007:235). Sebelum uji t dilanjutkan maka ada beberapa asumsi yang harus dipenuhi, yakni varian data harus homogen dan berdistribusi normal (Huck, 2012: 226).

#### **a. Uji homogenitas**

Uji homogenitas digunakan untuk membuktikan bahwa dua atau lebih kelompok sampel berasal dari populasi yang memiliki variansi yang sama. Pada penelitian ini uji homogenitas dilakukan dengan menggunakan tes *Levene* (Vorapongsathorn, Taejaroenkul, & Viwatwongkasem, 2004)

Hipotesis untuk menguji uji homogenitas ini adalah sebagai berikut:

$H_0 : \sigma_1^2 = \sigma_2^2$  (kedua kelompok populasi memiliki varians yang homogen)

$H_a : \sigma_1^2 \neq \sigma_2^2$  (kedua kelompok populasi memiliki varians yang tidak homogen)

Kriteria hasil analisis uji homogenitas dilakukan pada taraf signifikansi  $\alpha = 0,05$  (5%) adalah  $H_0$  ditolak jika nilai *p-value*  $< \alpha$  (Pallant, 2004: 234-235).

#### **b. Uji Normalitas**

Uji normalitas digunakan untuk membuktikan bahwa data sampel berasal dari populasi yang berdistribusi normal. Teknik yang digunakan untuk uji normalitas pada penelitian ini adalah dengan memperhatikan nilai *sig. Kolmogorov* dan *sig. Shapiro-Wilk* yang

merupakan tes uji normalitas yang memiliki kekuatan untuk semua jenis distribusi dan ukuran sampel (Razali & Wah, 2011).

Hipotesis untuk menguji uji normalitas ini adalah sebagai berikut:

$H_0$  : Data pada sampel berasal dari populasi yang berdistribusi normal

$H_a$  : Data pada sampel berasal dari populasi yang tidak berdistribusi normal

Analisis uji normalitas dilakukan dengan kriteria  $H_0$  ditolak jika  $p\text{-value} < 0,05$  (Zahedials & Ghasemi, 2012)

## 2. *Effect Size*

*Effect Size* merupakan uji yang dilakukan untuk mengetahui seberapa besar keefektifan perlakuan yang diberikan pada keterampilan berpikir tingkat tinggi siswa. Nilai *effect size* diperoleh dengan menggunakan rumus sebagai berikut (Cohen, 1988: 66-67)

$$d = \frac{\bar{X}_A - \bar{X}_B}{S_{pooled}}$$

$$S_{pooled} = \sqrt{\frac{\Sigma(X_A - \bar{X}_A)^2 + \Sigma(X_B - \bar{X}_B)^2}{n_A + n_B - 2}}$$

Keterangan:

$d$  = *Effect Size*

$\bar{X}_A$  = Rata-rata Kelas Eksperimen

$\bar{X}_B$  = Rata-rata Kelas Kontrol

$S_{pooled}$  = Standar Deviasi gabungan

$n_A$  = Jumlah sampel kelas eksperimen

$n_B$  = Jumlah sampel kelas kontrol

Kriteria untuk menginterpretasi hasil uji *effect size* mengikuti aturan seperti pada Tabel 16 (Cohen, 1988: 25-26)

**Tabel 16. Kriteria Nilai  $d$  (*Effect Size*)**

Nilai $d$	Kriteria
$0,8 < d < 2,0$	Kuat
$0,5 < d < 0,8$	Sedang
$0,2 < d < 0,5$	Lemah

### 3. Analisis Tingkat Kesukaran Soal

Analisis tingkat kesukaran soal adalah analisis yang bertujuan untuk mengetahui level kognitif yang berhasil dijawab dan sukar dijawab oleh siswa. Pada penelitian ini data dianalisis tingkat kesukarannya dengan menggunakan model *Rasch*. Tingkat kesukaran soal yang dianalisis dengan menggunakan model *Rasch* dikenal dengan sebutan model *one parameter logistic* (1PL). Model *Rasch* ini memprediksi kemungkinan keberhasilan setiap orang menjawab benar pada item  $i$  (Hambleton & Swaminathan, 1985: 69).

Model *Rasch* ini digambarkan sebagai model 1-PL karena mengandung satu parameter item, yakni tingkat kesukaran soal (Hambleton & Swaminathan, 1985: 69-70). Kemampuan seseorang dan perkiraan kesulitan butir soal, di komputer akan ditampilkan sepanjang skala logit. Item dan siswa terletak di peta yang sama. Skala logit adalah skala interval di mana semua unit logit memiliki ukuran yang sama. Nilai tertinggi terletak di bagian atas peta, dan nilai terendah terletak di bagian bawah. Setiap item dan siswa berada di

sepanjang skala logit sesuai dengan nilainya yang diperkirakan. Siswa yang berada pada logit positif (tinggi) merupakan siswa yang berkemampuan tinggi dan item soal tersebut merupakan soal yang sulit, sedangkan siswa yang berada pada logit yang rendah merupakan siswa dengan kemampuan rendah dan item soal tersebut merupakan soal dengan tingkat kesukaran rendah (Bond & Fox, 2007: 38-43).