

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **A. Metode Penelitian**

Penelitian ini merupakan sebuah penelitian deskriptif komparatif, karena dalam penelitian ini peneliti akan melakukan perbandingan dalam penetapan skor batas bawah dengan metode *Angoff* dan *Ebel*.

Pendekatan yang akan digunakan dalam penelitian ini adalah pendekatan kuantitatif. Melalui pendekatan ini peneliti akan melakukan perbandingan antara metode *Angoff* dan *Ebel* dengan menggunakan data keras. Menurut Ali (2011:416) data keras adalah bilangan atau angka-angka, baik diperoleh dari jumlah suatu penggabungan atau pun pengukuran. Data kuantitatif hasil pengukuran adalah skor-skor yang diperoleh melalui pengukuran, seperti angka hasil mengukur atau menimbang, skor skala rating dan skor jenis-jenis skala lainnya dan skor tes.

#### **B. Populasi dan Teknik Penarikan Sampel**

Populasi dalam penelitian ini adalah skor jawaban siswa kelas VII pada tes Bahasa Indonesia semester 1 kelas VII dan skor *judgement* panelis pada lembar kerja metode *Angoff* dan *Ebel* dengan responden siswa kelas VII pada tingkat MTs. dan guru kelas VII di MTs di Kabupaten Subang

Responden dalam penelitian ini adalah siswa kelas VII pada MTs Negeri Kasomalang dan MTs Cikaum di Kabupaten Subang, dengan responden panelis adalah guru kelas VII yang berada di lingkungan MTs Kabupaten Subang.

Secara geografis, Kabupaten Subang terbagi pada dua wilayah yakni Subang Utara dan Subang Selatan. Subang Utara meliputi daerah Tanjungsiang, Cislak, Kasomalang, Jalancagak, Ciater, dan Sagalaherang. Sedangkan Subang Selatan meliputi daerah Subang Kota, Pagaden, Kalijati, Cikaum, dan Pamanukan.

**Dewi Nuryawati, 2014**

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Adapun jumlah MTs di Subang Utara sebanyak 15 MTs dan MTs di Subang Selatan berjumlah 17 MTs (Data Seksi Madrasah Kandepag Subang) karena Kasomalang dan Cikaum merupakan daerah sentral pada masing-masing wilayah dengan data yang diperoleh berdasarkan hasil wawancara dengan guru dan kepala sekolah di MTs Negeri Kasomalang, siswa yang belajar disekolahnya tidak hanya berasal dari wilayah Kasomalang saja, tetapi ada juga siswa yang berasal dari Tanjungsiang, Cisalak dan Jalancagak. Begitupun dengan MTs Cikaum, berdasarkan hasil wawancara dengan guru dan kepala sekolah Cikaum, siswa tidak hanya berasal dari daerah Cikaum saja tetapi ada juga siswa yang berasal dari Pagaden, Binong dan Tambakdahan. Kemudian jumlah siswa kelas VII harus merupakan MTs yang mempunyai jumlah siswa kelas VII yang banyak sehingga diharapkan jumlah siswa yang banyak dapat mewakili siswa kelas VII MTs di Kabupaten Subang.

Penarikan sampel akan dilakukan dengan *simple random sampling* pada populasi penelitian dengan jumlah sampel pada setiap metode akan ditentukan sebesar 200 sampel yang terdiri dari 100 skor jawaban siswa dari MTs Negeri Kasomalang, dan 100 skor jawaban siswa dari MTs Cikaum. Sedangkan jumlah skor *judgement* panelis ditentukan dengan menggunakan persamaan Slovin. Berdasarkan daftar kehadiran dalam penetapan skor batas bawah yang dilaksanakan di MTs Negeri Kasomalang, jumlah guru yang menghadiri undangan berjumlah 37 orang. Adapun pengambilan jumlah sampel dilakukan dengan menggunakan persamaan Slovin, yakni :

$$n = \frac{N}{1+N \alpha^2} \quad (\text{Setiawan, 2007:6})$$

dimana,

n : banyaknya subjek penelitian

N : banyak populasi data penelitian

$\alpha$  : taraf signifikansi yang digunakan (0.05)

**Dewi Nuryawati, 2014**

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Peneliti menggunakan persamaan Slovin dalam penarikan sampel didasari pada pendapat Setiawan (2007) bahwa persamaan Slovin hanya digunakan jika :

1. Digunakan untuk menentukan ukuran sampel yang bertujuan menduga proporsi populasi
2. Asumsi tingkat keandalan 95% sehingga diperoleh nilai  $Z = 1,96$  yang kemudian dibulatkan menjadi  $Z = 2$
3. Asumsi keragaman populasi yang dimasukkan dalam perhitungan adalah  $P$  dimana  $P = 0,5$
4. Nilai galat pendugaan ( $d$ ) didasarkan atas pertimbangan peneliti.

Perhitungan dengan persamaan Slovin diperoleh hasil 33,86, maka jumlah sampel yang akan diambil berjumlah 34. Merujuk pada pendapat Roscoe dalam Sugiyono (2006:101) bahwa ukuran sampel yang layak adalah 30 – 500, dan diperkuat oleh Pett dan Salkind dalam Corder dan Foreman (2009:2) bahwa ukuran minimal agar dapat menggunakan statistik parametrik anggota sampel harus berjumlah  $n > 30$ , dan jumlah yang diperoleh dengan persamaan Slovin yakni 34 sampel telah memenuhi ukuran minimal jumlah sampel.

Penarikan sampel terpilih akan menggunakan *simple random sampling* dengan teknik perandoman sistematis. Menurut Ali (2011: 107) perandoman sistematis dilakukan dengan langkah awal yang harus dilakukan adalah membuat daftar pada seluruh unit populasi yang kemudian menentukan interval bilangan yang akan digunakan. Interval yang akan digunakan dalam penerikan sampel ini adalah interval 5 dan pada metode *Angoff* dimulai dari nomor urut 1, sedangkan pada metode *Ebel* dimulai dari nomor urut 2.

Jadi jumlah sampel yang akan digunakan pada penelitian ini adalah 200 skor jawaban siswa dan 34 skor *judgement* panelis. Jumlah ini sudah memenuhi jumlah minimal penerikan sampel, seperti Roscoe dalam Sugiyono (2006:101) bahwa ukuran sampel yang layak adalah 30 – 500.

**Dewi Nuryawati, 2014**

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

### C. Variabel Penelitian

Variabel bebas dalam penelitian ini adalah metode *Angoff* dan *Ebel*. Sedangkan variabel terikat adalah skor batas bawah pada mata pelajaran Bahasa Indonesia kelas VII di MTs Negeri Kasomalang dan MTs Cikaum. Berikut definisi operasional dari variabel terikat dan tidak terikat :

#### 1. Skor Batas Bawah

Skor batas bawah adalah skor minimal yang ditentukan oleh judgement guru dengan menggunakan metode *standard setting*. Skor batas bawah diperoleh dengan menjumlahkan semua judgement guru kemudian dicari rata-ratanya. Nilai rata-rata tersebut yang kemudian akan menjadi skor batas bawah. Selanjutnya skor tersebut dapat dijadikan skor minimal yang harus dicapai oleh siswa untuk memperoleh tingkat kelulusan suatu mata pelajaran tertentu dan dalam penelitian ini adalah mata pelajaran Bahasa Indonesia, dan yang akan diukur dalam penelitian ini adalah bagaimana pelaksanaan metode *Angoff* dan *Ebel* dalam menetapkan skor batas bawah, yakni dengan mengukur *standard error of judgement (SEj)*, dan *inter-judge*. Kemudian untuk memperjelas hasil penelitian dilakukan deskripsi data tentang pemetaan tingkat kesulitan butir dan jumlah ketuntasan siswa dengan menggunakan skor batas bawah yang diperoleh dengan metode *Angoff* dan *Ebel*.

#### 2. Metode *Angoff*

Metode *Angoff* adalah suatu metode *Standard setting* yang digunakan untuk menentukan skor batas bawah dengan merujuk pada skor *judgement* panelis yang telah dipilih. Pada metode ini panelis akan memperkirakan persentase jawaban benar siswa pada setiap butir tes. Skor batas bawah diperoleh dengan mencari rata-rata dari total skor *judgement* panelis dalam memperkirakan persentase jawaban benar siswa pada setiap butir tes. Berikut adalah prosedur pelaksanaan metode *Angoff* yang dilaksanakan dalam penelitian ini :

Dewi Nuryawati, 2014

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

1. Uji instrumen tes untuk memilih butir soal yang dapat mengukur kemampuan siswa dengan menggunakan program *ITEMAN* untuk melihat tingkat kesulitan butir soal dan tingkat daya beda butir soal.
2. Pemilihan panelis yang sesuai dengan karakteristik panelis yang telah ditentukan oleh peneliti
3. Uji coba pelaksanaan *judgement* panelis pada metode *Angoff*
4. Pelaksanaan metode *Angoff*
  - a. Setiap panelis diberikan sebuah instrumen tes Bahasa Indonesia Kelas VII yang sudah tervalidasi dan reliabel
  - b. Setiap panelis diberikan lembar kerja panelis metode *Angoff*
  - c. Setiap panelis secara individu melakukan estimasi berapa persen setiap butir soal akan dijawab benar oleh siswa
  - d. Lembar kerja panelis dikumpulkan dan dianalisis oleh peneliti

### **3. Metode *Ebel***

Metode *Ebel* adalah suatu metode *Standard setting* yang digunakan untuk menentukan skor batas bawah dengan merujuk pada skor *judgement* panelis yang telah dipilih. Pada metode ini panelis akan mengestimasi berapa persen butir soal akan dijawab benar oleh siswa berdasarkan pada setiap kategori. Ada 12 kategori dalam metode *Ebel*, yakni 3 kategori tingkat kesulitan (mudah, sedang, sukar) dan 4 kategori relevansi isi (esensial, penting, dapat diterima, dapat dipertanyakan). Berikut prosedur pelaksanaan metode *Ebel* yang dilaksanakan dalam penelitian ini, yakni :

1. Uji instrumen tes untuk memilih butir soal yang dapat mengukur kemampuan siswa dengan menggunakan program *ITEMAN* untuk melihat tingkat kesulitan butir soal dan tingkat daya beda butir soal.
2. Pemilihan panelis yang sesuai dengan karakteristik panelis yang telah ditentukan oleh peneliti
3. Uji coba pelaksanaan *judgement* panelis pada metode *Ebel*

Dewi Nuryawati, 2014

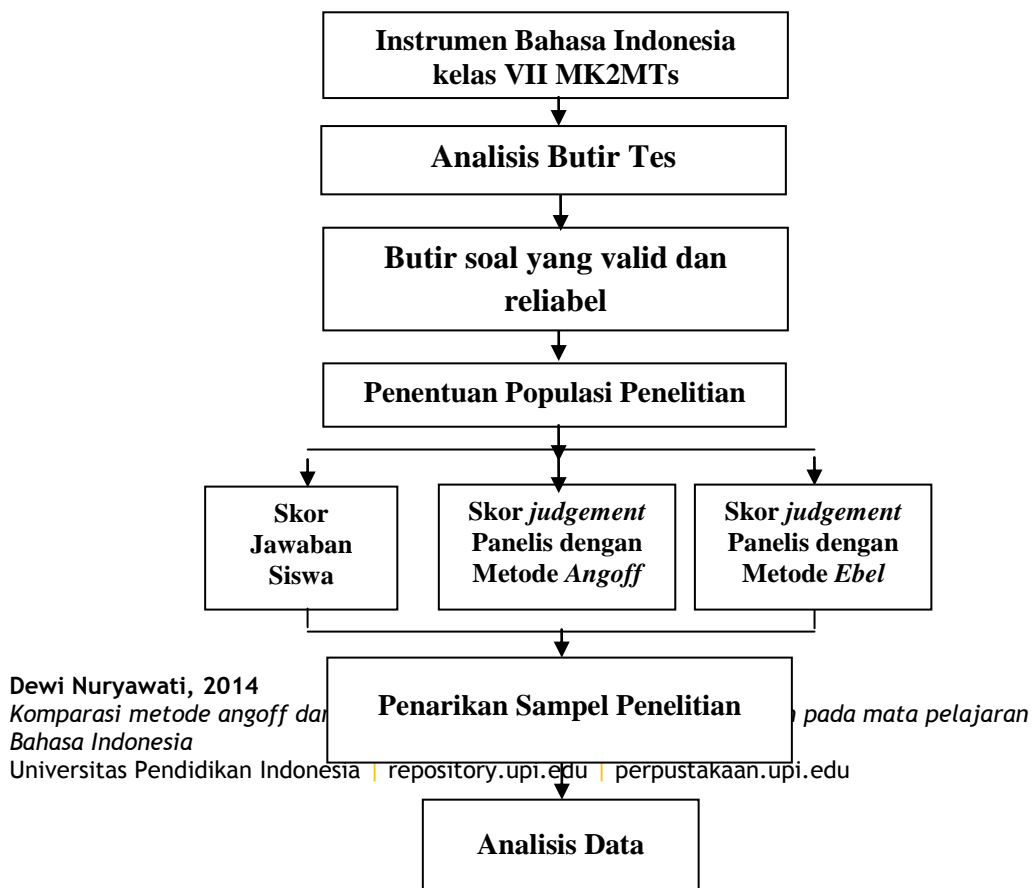
*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

#### 4. Pelaksanaan metode *Ebel*

- a. Setiap panelis diberikan sebuah instrumen tes Bahasa Indonesia Kelas VII yang sudah tervalidasi dan reliabel
- b. Setiap panelis diberikan lembar kerja panelis metode *Ebel*
- c. Setiap panelis secara individu menentukan setiap butir soal pada 12 kategori dalam metode *Ebel*
- d. Mendiskusikan kembali secara kelompok setiap butir soal pada 12 kategori dalam metode *Ebel*
- e. Setiap panelis secara individu melakukan estimasi berapa persen setiap kategori butir soal akan dijawab benar oleh siswa
- f. Lembar kerja panelis dikumpulkan dan dianalisis oleh peneliti

#### D. Alur Penelitian



### **Gambar : 3.1 Alur Penelitian**

#### **1. Analisis Butir Tes Bahasa Indonesia Kelas VII**

Tes pada penelitian ini menggunakan tes Bahasa Indonesia yang digunakan siswa Kelas VII di Mts pada Ujian Akhir Semester (UAS) Satu Tahun Ajaran 2013-2014. Butir tes dibuat oleh MK2MTs yakni sebuah kelompok kerja MTs Provinsi Jawa Barat. Namun karena pada penelitian ini harus menggunakan sebuah tes yang valid dan reliabel, maka uji validitas dan reliabilitas dilakukan sebelum diberikan pada panelis.

Uji validitas isi oleh tim terhadap 40 butir tes pilihan ganda yang terdapat pada instrumen MK2MTs, dengan menggunakan uji validitas isi persentase butir yang cocok dengan indikator dan tujuan. Kemudian dilakukan uji tingkat kesulitan dan daya beda butir untuk mengetahui butir mana saja yang digunakan sebagai butir dalam instrumen dan butir mana saja yang akan dibuang, dan dilakukan juga uji reliabilitas tes untuk mengetahui berapa kuat sebuah instrumen dapat digunakan sehingga responden membuat judgement pada butir tes yang tepat. Uji Tingkat kesulitan butir, daya beda dan reliabilitas dilakukan dengan bantuan *Iteman*.

#### **2. Butir Tes yang Valid dan Reliabel**

Setelah diperoleh butir tes yang valid dan reliabel, maka tes akan diberikan kepada panelis yang telah terpilih sebagai rujukan untuk membuat sebuah *judgement* pada metode *Angoff* dan *Ebel*.

#### **3. Penentuan Populasi Penelitian**

Dewi Nuryawati, 2014

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Populasi penelitian akan ditentukan oleh peneliti dengan pertimbangan populasi dapat mewakili skor jawaban siswa dan skor *judgement* panelis pada metode *Angoff* dan *Ebel* di kelas VII MTs di Kabupaten Subang.

#### **4. Penarikan Sampel Penelitian**

Penarikan sampel akan dilakukan dengan teknik *simple random sampling* pada kedua populasi penelitian yakni populasi skor jawaban siswa dan skor *judgement* panelis pada metode *Angoff* dan *Ebel*.

#### **5. Analisis Data.**

Analisis dilakukan berdasarkan pada data yang telah diperoleh dengan pembahasan yang telah didefinisikan dalam definisi operasional, yakni melakukan komparasi antara metode *Angoff* dan *Ebel*.

#### **E. Teknik Pengumpulan Data**

Instrumen yang digunakan dalam pengumpulan data adalah instrumen tes siswa dan lembar kerja panelis dengan menggunakan metode *Angoff* dan *Ebel*. Instrumen tes siswa digunakan untuk memperoleh lembar jawab, dan lembar kerja panelis digunakan untuk memperoleh skor *judgement* panelis.

#### **F. Teknik Analisis Data**

##### **1. Uji Instrumen Penelitian**

Instrumen tes akan merujuk pada instrumen tes Bahasa Indonesia kelas VII yang dibuat oleh MK2MTs yang digunakan sebagai tes UAS satu Tahun Ajaran 2013 - 2014. Namun, sebelum dijadikan tes yang dijadikan rujukan dalam membuat *judgement* dikelas penetapan skor batas bawah, butir tes akan dianalisis terlebih dahulu untuk mengetahui tingkat validitas dan reliabilitas tes. Uji coba instrumen dilakukan pada siswa Kelas VII di MTs Negeri Kasomalang dan MTs Cikaum pada hasil jawaban siswa pada UAS semester 1. Kedua sekolah tersebut dipilih dengan pertimbangan dan rekomendasi Seksi Madrasah Kandepag Subang

Dewi Nuryawati, 2014

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu



bahwa jumlah siswa pada kedua MTs tersebut mewakili siswa MTs Kelas VII di Kabupaten Subang karena jumlah siswa kelas VII pada kedua MTs tersebut merupakan siswa kelas VII yang terbanyak, yakni siswa MTs Negeri Kasomalang yang berjumlah 290 orang dan siswa MTs Cikaum yang berjumlah 197 orang.

Agar pelaksanaan penelitian berjalan sesuai dengan tujuan penelitian, maka instrumen yang akan diberikan pada responden adalah kumpulan dari butir tes Bahasa Indonesia yang valid serta instrumen yang reliabel, maka peneliti melakukan beberapa analisis tes, yakni :

#### **a. Validitas Isi**

Validitas isi dilakukan untuk mengetahui kesesuaian butir tes dengan indikator atau materi pembelajaran yang telah ditetapkan. Butir tes dinyatakan valid, jika butir-butir yang dibuat sesuai dengan indikator atau materi-materi yang telah ditetapkan (Susetyo, 2011).

Teknik analisis validitas isi dalam penelitian ini menggunakan persentase butir yang cocok dengan indikator atau tujuan karena instrumen ini menggunakan skala dikotomi yaitu skor 1 untuk jawaban yang cocok dan skor 0 untuk jawaban yang tidak cocok. Pada teknik ini, butir tes dinyatakan valid jika ada kecocokan antara indikator dengan butir tes mencapai lebih besar dari 50%.

Persamaan yang digunakan adalah :

$$\text{Persentase} = \frac{f}{\sum f} \times 100\% \quad (\text{Susetyo, 2011:92})$$

dimana:

F = frekuensi cocok menurut penilai

$\sum f$  = jumlah penilai

**Dewi Nuryawati, 2014**

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

dan berdasarkan teknik validitas kecocokan antara indikator dengan butir tes ada 2 butir tes yakni butir tes nomor 9 dan 20 yang dinyatakan tidak valid dan 38 butir tes yang dinyatakan valid.

### b. Tingkat Daya Beda

Uji daya beda tes pun dilakukan, uji ini dilakukan untuk mengukur seberapa besar suatu butir tes dapat membedakan antar siswa yang mempunyai kemampuan tinggi dan siswa yang mempunyai kemampuan rendah.

Daya beda dapat dilihat dari besar kecilnya indek diskriminasi, menurut Nunnally (Susetyo, 2012:161) koefisien korelasi di atas 0.2 sudah dianggap cukup baik. Sedangkan koefisien korelasi yang berada di bawah 0.2, butir tes harus dibuang karena butir tersebut kurang berada dalam satu kesatuan perangkat tes dengan butir tes lainnya. Uji daya beda dalam penelitian ini menggunakan bantuan program *Iteman*, dan skor daya beda akan ditunjukkan pada kolom *Point Biserial*.

Setelah dilakukan uji instrumen hanya 23 butir tes yang memiliki daya beda diatas 0.2 dan 27 butir harus dibuang. Berikut adalah pembagian daya beda menurut Ebel dalam Susetyo (2012:161) :

**Tabel 3.1 Kriteria Indek Daya Beda**

<b>Indek Daya Beda</b>	<b>Keterangan</b>
$0,70 \leq D \leq 1,0$	Butir memiliki daya beda baik sekali
$0,40 \leq D \leq 0,69$	Butir memiliki daya beda cukup baik
$0,30 \leq D \leq 0,39$	Butir memerlukan revisi sedikit atau tidak
$0,20 \leq D \leq 0,29$	Butir memerlukan revisi atau disisihkan
$0,00 \leq D \leq 0,19$	Butir direvisi total atau disisihkan

### c. Reliabilitas Tes

Dewi Nuryawati, 2014

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Reliabilitas tes menjadi salah satu hal yang harus dilakukan untuk mengukur tingkat keajegan suatu tes. Suatu perangkat tes dinyatakan reliabel jika telah mencapai sekurang-kurangnya memperoleh koefisien korelasi sebesar 0,50 (Susetyo, 2011:107). Peneliti akan menggunakan KR<sub>20</sub> untuk menguji tingkat reliabilitas tes karena hasil tes akan diberikan bobot skor dikotomi yakni 1 untuk jawaban yang benar dan 0 untuk jawaban yang salah. Perhitungan dilakukan secara langsung pada seluruh butir tes. KR<sub>20</sub> akan menunjukkan kekurangsetaraan semua butir tes mengakibatkan terjadinya interkorelasi diantara butir menjadi rendah. Persamaan dapat dilihat dibawah ini:

Persamaan KR<sub>20</sub>

$$\rho_{KR20} = \frac{k}{k-1} \left( 1 - \frac{\sum pq}{\sigma_A^2} \right) \quad (\text{Susetyo,2011:116})$$

dimana,

- P = proporsi menjawab benar
- Q = proporsi menjawab salah
- K = jumlah butir tes
- $\sum pq$  = jumlah perkalian jawaban benar dengan salah
- $\rho_{KR20}$  = koefisien reliabilitas
- $\sigma_A^2$  = varian skor tes
- N = jumlah responden

dan hasil pengukuran dengan KR<sub>20</sub> tingkat reliabilitas instrumen adalah 0.754, karena skor yang diperoleh lebih dari 0.5 maka instrumen dinyatakan reliabel.

#### **d. Tingkat Kesulitan Butir Tes**

Tingkat kesulitan merupakan perbandingan antara siswa yang menjawab benar dengan jumlah butir tes. Tingkat kesulitan menggambarkan seberapa sukar suatu butir tes dijawab oleh responden. Untuk mengetahuinya, peneliti melakukan perhitungan atas hasil jawaban siswa dengan bantuan program *Iteman*. Tingkat kesulitan butir tes dilihat pada kolom hasil *Prop. Correction*. Menurut

Dewi Nuryawati, 2014

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Witherington (Susetyo, 2012:154) pembagian tingkat kesulitan adalah sebagai berikut :

**Tabel 3.2 Kriteria Tingkat Kesulitan**

Rentang	Tingkat Kesulitan
$0,00 \leq P \leq 0,24$	Sukar
$0,25 \leq P \leq 0,74$	Sedang
$0,75 \leq P \leq 1,00$	Mudah

Kemampuan peserta tampak tinggi apabila tingkat kesulitan butirnya tergolong mudah. Sebaliknya kemampuan tampak rendah, apabila tingkat kesulitan butir tergolong sukar. Kesulitan butir tampak mudah, apabila kemampuan peserta tergolong tinggi. Sebaliknya kemampuan peserta rendah, apabila tingkat kesulitan butir tergolong tinggi.

Berdasarkan hasil perhitungan dengan bantuan *Iteman* , 23 tes yang mempunyai daya beda  $\Rightarrow$  0,20 dan 6 butir tes yang tergolong kategori mudah, 15 butir tes sedang dan 2 butir tes sukar.

## 2. Komparasi

Seperti definisi operasional dalam penelitian ini, komparasi yang akan dilakukan dengan membandingkan metode *Angoff* dan *Ebel* dalam menetapkan skor batas bawah dengan cara menganalisis validitas metode dengan mengukur konsistensi metode dan konsistensi *inter-judge* metode,

### a. Konsistensi Metode

Konsistensi metode dilakukan untuk mengukur *standard error of judgement* (SEj), hal ini dilakukan untuk mengukur apakah metode yang akan digunakan memadai atau tidak jika digunakan sebagai metode yang digunakan untuk menetapkan skor batas bawah. pengukuran dilakukan dengan melihat SEj, dan konsistensi dikatakan memadai jika nilai SEj sama dengan atau lebih kecil dari setengah nilai *standard error of measurement* (SEM) tes. Menurut Norcini dalam Report (2012:8), SEj diperoleh dengan cara membagi standar deviasi dari *judgement* panelis dengan *square root* dari banyak panelis.

Dewi Nuryawati, 2014

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

### **b. Konsistensi *Inter-judge***

Pengukuran tentang *inter-judge* dilakukan untuk mengetahui kelayakan suatu metode, pada penelitian ini untuk mengukur apakah metode *Angoff* dan *Ebel* merupakan metode yang layak sebagai metode untuk menetapkan skor batas bawah. pengukuran dilakukan dengan melihat simpangan baku skor batas bawah pada metode *Angoff* dan *Ebel*. Semakin kecil simpangan baku yang diperoleh suatu metode, maka semakin layak metode digunakan untuk menetapkan skor batas bawah.

### **c. Deskripsi Perbandingan**

Selain mengukur validitas metode, peneliti juga melakukan deskripsi perbandingan antara metode *Angoff* dan *Ebel*. Deskripsi perbandingan ini dilakukan dengan cara membandingkan pemetaan tingkat kesulitan yang dilakukan dengan program *ITEMAN* dengan skor *judgement* panelis, kemudian juga melihat jumlah ketuntasan siswa pada sampel skor jawaban siswa kelas VII MTs berdasarkan skor batas bawah yang diperoleh dengan metode *Angoff* dan *Ebel*.

## **3. Uji Hipotesis**

Setelah perbandingan validitas dan perbandingan secara deskriptif dilakukan, maka akan dilakukan uji hipotesis, untuk mengetahui apakah terdapat perbedaan yang signifikan dalam metode *Angoff* dan *Ebel* dalam menetapkan skor batas bawah. uji hipotesis dilakukan dengan analisis inferensial. Analisis ini dilakukan karena dalam penelitian ini akan menaksir parameter populasi berdasarkan ukuran-ukuran sampel dan menguji hipotesis. Melalui sampel yang terpilih, dengan statistik inferensial peneliti membuat kesimpulan-kesimpulan tentang parameter populasi (Furqon, 2009:145). Berikut adalah langkah-langkah dalam uji hipotesis penelitian :

**Dewi Nuryawati, 2014**

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

1. Panelis menganalisis instrumen tes Bahasa Indonesia.
2. Panelis membuat *judgement* dalam lembar kerja.
3. Menghitung *judgement* panelis
4. Menghitung skor batas bawah berdasarkan jumlah data dalam penetapan skor batas bawah dengan metode *Angoff* dan *Ebel*.
5. Uji persyaratan normalitas dan uji homogenitas. Uji normalitas akan dilakukan melalui *uji Kolmogorov Smirnov* dan uji homogenitas dengan *Uji Lavene*
6. Uji statistik yang akan digunakan adalah uji t untuk melakukan uji signifikansi perbedaan antara dua rata-rata dari dua sampel, karena dalam penelitian ini skor yang diperoleh berasal dari sampel yang sama, maka uji perbedaan menggunakan *dependent T test*.

#### 4. Hipotesis Statistik

Hipotesis statistik dalam penelitian ini adalah sebagai berikut :

1. Skor batas bawah pada metode *Angoff* tidak berbeda dengan skor batas bawah pada metode *Ebel* ( $H_0$ ) ;  $H_0: \rho_1 = \rho_2$
2. Skor batas bawah pada metode *Angoff* berbeda dengan skor batas bawah pada metode *Ebel* ( $H_1$ ) ;  $H_1: \rho_1 \neq \rho_2$

Dewi Nuryawati, 2014

*Komparasi metode angoff dan ebel dalam menetapkan skor batas bawah pada mata pelajaran Bahasa Indonesia*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu