

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **1.1 KAJIAN TEORI**

Beberapa teori yang digunakan dalam penelitian ini yaitu teori kualitas, pengertian instrumen, pengertian penilaian, penyusunan perangkat tes, analisis butir soal.

##### **1.1.1 Pengertian penilaian**

Penilaian adalah suatu proses mengukur untuk menentukan pencapaian tujuan pembelajaran. Tuckman dalam (Nurgiyantoro, 2016: 6) mendefinisikan penilaian sebagai suatu proses untuk mengetahui atau menguji suatu kegiatan, keluaran suatu program, dan proses apakah sesuai dengan tujuan atau kriteria yang telah ditentukan.

Brown menganggap bahwa penilaian sama dengan tes, menurut Brown penilaian adalah pengukuran pengetahuan, kemampuan, dan kinerja seseorang dalam suatu ranah yang diberikan. Dalam definisi tersebut Brown menganggap bahwa penilaian paling tidak mengandung tiga hal. Pertama, tes sebagai sebuah cara: tes merupakan sejumlah teknik, prosedur, atau butir yang menuntut jawaban atau kinerja seseorang. Kedua, tes sebagai alat ukur: tes mesti mengukur sesuatu dalam diri seseorang baik yang berupa pengetahuan umum maupun kemampuan khusus. Ketiga, tes sebagai alat ukur pengetahuan, kemampuan, atau kinerja: tes dimaksudkan untuk mengukur sesuatu atau ciri laten seseorang seperti pengetahuan, kemampuan, keterampilan, kinerja, dan lain-lain (Nurgiyantoro, 2016:9-10).

Dari pengertian tersebut dapat disimpulkan bahwa penilaian adalah pengukuran hasil kegiatan yang telah dilaksanakan oleh guru dan siswa dalam suatu kegiatan pembelajaran untuk mengetahui (mengukur) apakah tujuan pembelajaran tercapai atau tidak. Penilaian juga dapat diartikan suatu proses pengumpulan informasi yang dipergunakan sebagai dasar pembuatan keputusan tentang program pendidikan, jadi penilaian memiliki komponen pengumpulan informasi, penggunaan informasi, dan pembuatan keputusan (Cronbarch dalam Nurgiantoro, 2016:188)

Informasi yang dimaksud adalah informasi tentang pihak yang akan dinilai dalam pembelajaran atau peserta didik. Informasi berupa pengetahuan, kemampuan, keterampilan, kinerja, dan lainnya yang diperoleh peserta didik selama kegiatan pembelajaran berdasarkan kompetensi yang telah dipelajari. Pemerolehan informasi yang akurat juga akan menyebabkan pertimbangan yang dilakukan itu tepat. Jika informasi yang didapat tidak akurat maka pertimbangan yang dilakukan tidak akan

tepat sesuai dengan tujuan, maka dari itu pembuatan pertimbangan harus tepat sesuai dengan informasi yang di dapatkan.

Pembuatan pertimbangan adalah hal yang penting serta harus dilakukan dalam kegiatan penilaian. Agar informasi yang didapatkan sesuai dengan kebutuhan, pengumpulan informasi berdasarkan pada rencana pertimbangan yang akan dibuat. Pertimbangan merupakan tafsiran terhadap kondisi yang ada pada sekarang dan prediksi untuk masa mendatang.

Tujuan akhir yang dilakukan dalam kegiatan penilaian adalah pengambilan keputusan. Pembuatan keputusan adalah pemilihan diantara sejumlah alternatif. Pembuatan keputusan tidak harus orang yang membuat yang membuat pertimbangan tetapi boleh juga orang yang sama.

Ada keterkaitan antara pembelajaran (teaching), belajar (learning), dan penilaian (evaluation). Menurut Gronlund (Nurgiantoro 2016:14) proses pelaksanaan penilaian dan pembelajaran. Proses yang dimaksud adalah proses pembelajaran difasilitasi oleh guru agar peserta didik dapat belajar secara maksimal sesuai dengan kompetensi yang diajarkan.

### **1.1.2 Hakikat instrument**

Instrumen penilaian harus memenuhi persyaratan substansi yang merepresentasikan kompetensi yang dinilai, konstruksi yang memenuhi persyaratan teknis sesuai dengan bentuk instrumen yang digunakan, dan penggunaan bahasa yang baik dan benar serta komunikatif sesuai dengan tingkat perkembangan peserta didik (Kunandar dalam Pangesti, 2016:17).

Menurut Mardapi ( dalam Nurgiyantoro 2016:20) Pengembangan instumen penilaian mengikuti langkah-langkah sebagai berikut :

#### **1) Penentuan spesifikasi ujian**

Dalam penentuan spesifikasi ujian paling tidak terdapat empat hal yang harus dilakukan yaitu:

##### **a. penentuan kompetensi dasar**

pengujian dimaksudkan untuk untuk mengukur seberapa banyak peserta didik dapat mencapai kompetensi yang dipelajari. Kompetensi yang diukur kadar capainya adalah kompetensi dasar yang akan dijabarkan dari standart kompetensi. Standart kompetensi dan kompetensi dasar yang dikhususkan pada mata pelajaran bahasa Indonesia dikaitkan dengan empat keterampilan berbahasa yakni menyimak, berbicara, membaca, dan menulis.

b. Pembuatan deskripsi bahan uji

Jika kompetensi dasar yang akan diambil tujuannya dengan jelas maka akan mudah untuk meraih kompetensi yang dimaksud. Pembuatan deskripsi bahan ajar yang meliputi materi pokok dan uraian materi haruslah dilakukan untuk memastikan bahan ajar apa saja yang akan diujikan.

c. Pembuatan kisi-kisi pengujian

Agar kegiatan pengembangan dapat dilakukan dengan baik dan dapat dipertanggungjawabkan, pembuatan butir-butir soal (pertanyaan) harus berdasarkan pada kisi-kisi yang telah dibuat.

d. Penentuan bentuk soal dan lama ujian

Penentuan bentuk soal untuk ujian akhir semester biasanya berbentuk pilihan ganda, selain itu juga harus ditentukan berapa soal yang dibuat berdasarkan waktu lama ujian itu berlangsung

2) Penulisan butir soal

Penulisan butir soal adalah membuat tagihan sesuai dengan tuntutan indicator dan bahan ajar. Butir soal harus cocok dengan kata kerja operasionalnya. Penulisan butir-butir soal juga harus memperhatikan langkah-langkah yang digunakan untuk telaah butir soal agar tidak banyak revisi.

3) Penelaahan butir soal

Untuk memastikan butir-butir soal yang ditulis telah memenuhi tuntutan soal yang baik maka sebelum diujicobakan harus terlebih dahulu dilakukan telaah butir soal. Criteria penelaahan butir soal dapat dilihat dari materi, konstruksi dan bahasa yang digunakan.

4) Pelaksanaan uji coba

Jika semua persyaratan penyusunan butir soal sudah dilakukan maka tahap selanjutnya adalah pelaksanaan uji coba untuk mengetahui instrument yang telah dibuat baik dan dapat dipertanggungjawabkan.

5) Analisis butir soal dan jawaban

Analisis butir soal dapat menggunakan pengukuran klasik digunakan untuk menghitung indeks tingkat kesukaran (ITK) yang akan menentukan seberapa mudah atau sulitnya sebuah soal, indeks daya beda (IDB) yang membedakan daya butir soal termasuk tinggi atau rendah, dan efektifitas distraktor tentang bagaimana kemampuan *distraktor* untuk mengecoh peserta ujian.

Analisis jawaban adalah telaah hasil pengukuran per indicator per kemampuan dasar yang telah dilakukan dengan menghitung jawaban benar dan salah peserta didik untuk seluruh butir soal yang telah diujikan.

6) Perbaiki butir soal dan perakitan soal ujian

Hasil analisis butir soal dan jawaban pasti banyak informasi yang akan didapatkan. Pasti terdapat butir –butir soal yang harus diperbaiki baik secara tulisan maupun keakuratan.

7) Pelaksanaan ujian

Pelaksanaan ujian harus dilakukan sebaik mungkin dengan pengawasan yang cermat, tetapi tidak menimbulkan tekanan pada peserta didik

8) Penafsiran hasil ujian

Pelaksanaan pengujian akan menghasilkan hasil empiric kuantitatif yang berwujud skor-skor untuk tiap peserta didik yang kemudian ditafsirkan untuk memberikan makna capaian peserta didik.

### 1.1.3 Tes Pilihan Ganda

Tes pilihan ganda banyak digunakan dalam dunia pendidikan tes pilihan ganda merupakan tes yang terdiri dari beberapa pertanyaan yang belum lengkap dan masing-masing pertanyaan a terdapat pilihan jawaban yang benar dan jawaban untuk mengecoh. Pembahasan dalam tes pilihan ganda terdapat 5 pokok komponen yaitu kelebihan tes pilihan ganda, kelemahan tes pilihan ganda, ciri-ciri pokok, pengembangan, dan penggunaan tes pilihan ganda.

Pembahasan yang pertama yaitu kelebihan tes pilihan ganda. Tes pilihan ganda memiliki sejumlah kelebihan. Kelebihan tes pilihan ganda yaitu (1) peluang untuk jawaban benar dengan sekadar menebak dibandingkan tes benar salah, (2) cakupan materi tes yang luas, (3) cara menjawab yang sederhana, (4) pemeriksaan jawaban yang lebih sederhana, (5) analisis yang lebih mudah dilakukan terhadap masing-masing butir tes maupun tes secara keseluruhan karena sekadar didasarkan atas jumlah atau persentasi, termasuk penghitungan reliabilitas (reliability) tes (Djiwandono dalam pangesti, 2016:19).

ciri-ciri pokok tes pilihan ganda yaitu sejenis tes objektif yang masing-masing butir tesnya terdapat lebih dari dua pilihan jawaban atau pengecoh. Satu butir tes pilihan ganda terdapat pernyataan pokok atau pertanyaan yang selanjutnya diikuti oleh beberapa pilihan jawaban yang salah satunya benar atau merupakan kunci

jawaban dan yang lainnya merupakan pengecoh atau pilihan jawaban yang salah . jumlah pilihan yang paling umum digunakan paling tidak ada 3 atau lebih, masing-masing pilihan diberi tanda biasanya dengan tanda (a), (b), (c), dan (d). dari semua pilihan tersebut yang benar merupakan jawaban yang sering disebut jawaban kunci. Sedangkan pilihan yang lain merupakan jawaban yang tidak benar dibandingkan dengan jawaban benar yang sesungguhnya. Pilihan di luar jawaban benar disebut pengecoh (Djiwandono dalam pangesti, 2016:19).

Supranata (2005:179-193) Kaidah-kaidah dalam penulisan soal merupakan pedoman yang harus diikuti penyusun agar soal yang dibuat menghasilkan mutu yang baik. Soal yang mutunya baik ialah soal yang mampu memberi informasi berdasarkan tujuan yang diperlukan dan berfungsi selayaknya. Kualitas suatu tes sebenarnya banyak ditentukan oleh kualitas tiap butir soal yang disusun. Oleh karena itu dalam penulisan soal hendaknya memperhatikan kaidah dalam penulisan soal yang baik.

Salah satu penyusunan yang disarankan oleh Nurgiyantoro (2016:150) adalah sebagai berikut:

- 1) Pernyataan pokok hendaknya berisi satu permasalahan.
- 2) Pada setiap butir soal hanya terdapat satu jawaban yang tepat. Alternatif lainnya berlaku sebagai pengecoh dan harus menunjukkan unsure tertentu yang memang salah. Hal yang harus dihindari adalah adanya butir pengecoh yang mempunyai kemungkinan benar sehingga akan menjadi perdebatan dengan jawaban yang benar.
- 3) Pilihan jawaban yang disediakan harus mempunyai hubungan gramatikal yang benar dan sesuai dengan pertanyaan.
- 4) Panjang tiap alternatif jawaban hendaklah sama.
- 5) Harus menghindari pemberitahuan yang benar secara tidak langsung dan mungkin terlihat pada butir soal selanjutnya. Oleh karena itu soal yang satu dengan lainnya hendaknya tidak berkaitan.
- 6) Menghindari jawaban benar dengan berpola tertentu.

#### **1.1.4 Indeks tingkat kesukaran (ITK)**

Tingkat kesukaran merupakan pernyataan tentang seberapa mudah atau sulit butir soal bagi peserta didik yang dikenai pengukuran (Oller dalam Nurgiyantoro : 218). Oller juga menjelaskan seberapa butir soal memberi kemudahan kepada peserta

didik. Butir soal yang baik adalah butir soal yang tingkat kesukarannya tidak terlalu mudah dan tidak terlalu sulit.

Menganalisis tingkat kesukaran soal artinya mengkaji soal-soal tes dari segi kesulitan pada soal sehingga dapat dilihat soal-soal yang termasuk dalam kategori sedang, dan sukar. Terdapat beberapa pertimbangan untuk menentukan proporsi jumlah kategori mudah, sedang, dan sukar. Pertimbangan yang pertama adalah terdapat adanya keseimbangan, yaitu jumlah soal yang sama untuk masing-masing kategori tersebut. Artinya jumlah soal mudah, sedang, dan sukar harus seimbang. Pertimbangan kedua proporsi jumlah soal untuk ketiga kategori tersebut berdasarkan pada kurva normal. Artinya, sebagian besar soal berada dalam kategori sedang, sebagian lagi termasuk ke dalam kategori mudah dan sukar dengan proporsi yang seimbang.

Cara melakukan analisis tingkat kesukaran dengan rumus sebagai berikut:

$$ITK = \frac{B}{N}$$

ITK = Indeks tingkat kesulitan

B = Jumlah siswa dengan jawaban benar

N = Banyaknya siswa yang menjawab soal

### 1.1.5 Indeks daya beda (IDB)

Daya beda pada butir soal merupakan suatu pernyataan tentang seberapa besar butir soal dapat membedakan kemampuan antara peserta kelompok tinggi dan peserta dengan kelompok rendah. Indeks daya beda adalah indeks yang menunjukkan seberapa besar butir soal kemampuan antara peserta kelompok tinggi dan kelompok rendah (Nurgiyantoro: 221)

Dari definisi diatas dapat dijelaskan bahwa peserta uji dengan kelompok tinggi akan mampu menjawab dengan benar butir-butir soal daripada yang dikerjakan oleh kelompok rendah yang akan menjawab dengan salah. Jika kelompok rendah menjawab benar dengan jumlah lebih banyak daripada kelompok tinggi maka soal tersebut perlu dipertanyakan atau dalam kategori tidak baik. Karena soal yang baik adalah soal yang mempunyai daya untuk membedakan kelompok yang mempunyai kemampuan tinggi dan kelompok yang rendah.

Untuk menghitung daya beda pada soal dapat dilakukan dengan menggunakan rumus berikut :

$$IDB = \frac{FTK-FKR}{N}$$

IDB = Indeks daya beda

FTK = Jumlah jawaban benar kelompok tinggi

FKR = Jumlah jawaban benar kelompok rendah

N = jumlah peserta kelompok tinggi dan rendah

### 1.1.6 Pengecoh pilihan jawaban (*Distraktor*)

Dalam soal pilihan ganda terdapat pilihan jawaban yang disediakan, dan pilihan jawaban tersebut harus bias mempunyai peran untuk mengecoh peserta uji. Artinya walau opsi yang dibuat itu salah maka opsi-opsi tersebut tetap harus ada sejumlah peserta yang memilih. Opsi salah harus memiliki peran untuk menjebak, merusak atau pengecoh yang baik terhadap sebagian peserta uji.

Beberapa criteria distraktor yang efektif untuk menurut Nurgiyantoro adalah sebagai berikut:

- a. Semua pengecoh opsi salah harus ada yang memilih
- b. Jumlah pemilih opsi dari peserta kelompok tinggi harus lebih sedikit daripada kelompok rendah
- c. Jika pemilih opsi salah hanya satu maka harus dari kelompok rendah

## 1.2 PENELITIAN TERDAHULU YANG RELEVAN

Peneliti mencatat terdapat dua penelitian yang relevan. Penelitian pertama dilakukan oleh Viviyanti Dyah Pangesti (2016) dengan judul “*Pengembangan Instrumen Penilaian Kompetensi Dasar Menyimak Pada Pembelajaran Bahasa Indonesia Untuk Siswa Kelas VII SMP N 15 Yogyakarta*”. Penelitian ini menemukan beberapa hal yaitu proses pengembangan untuk menghasilkan instrument penilaian dasar menyimak untuk kelas VII meliputi (1) analisis kebutuhan, (2) menyusun spesifikasi produk instrument penilaian kemampuan dasar menyimak (3) menyusun instrument penilaian kemampuan dasar menyimak (4) menelaah instrument penilaian (5) melakukan revisi satu berdasarkan hasil penilaian ahli penelitian (6) melakukan uji coba produk instrument penilaian kepada peserta didik (7) melakukan analisis hasil uji coba produk instrumen penilaian (8) melakukan revisi terhadap produk hasil uji coba siswa (8) melakukan revisi terhadap produk instrumen penilaian berdasarkan hasil uji coba pada peserta didik (9) menghasilkan produk pengembangan instrument penilaian kemampuan dasar menyimak.

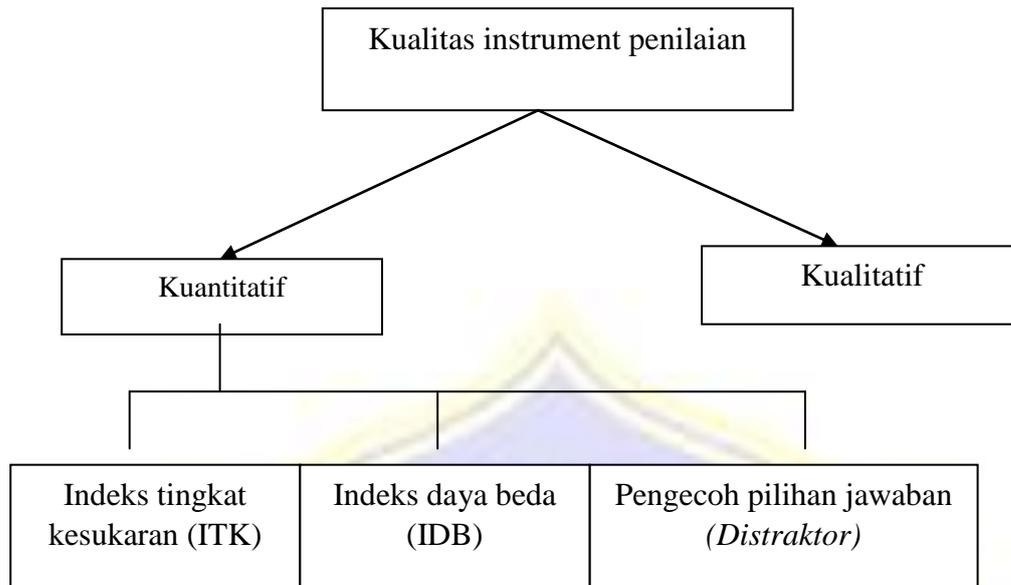
Relevansi penelitian ini dengan penelitian yang sedang dilakukan adalah sama-sama meneliti tentang instrument penilaian namun yang diteliti berbeda dimana penelitian oleh Vividyah adalah untuk mengembangkan instrument penilaian keterampilan dasar menyimak sedangkan yang sedang diteliti saat ini adalah untuk mengetahui kualitas instrument penilaian mata pelajaran bahasa Indonesia di sekolah. Karena sama-sama meneliti tentang instrument penilaian maka terdapat referensi untuk melakukan penelitian ini.

Penelitian yang kedua penelitian yang dilakukan oleh A. Eris Eriyansah, 2013 dengan judul "*Evaluasi Alat Penilaian Ujian Kenaikan Kelas Mata Pelajaran Produktif di SMK Negeri 1 Karangtengah Cianjur*" penelitian ini menjelaskan tentang evaluasi alat yang digunakan dalam penilaian ujian kenaikan kelas yang dikaji dengan kuantitatif dan hasilnya terdapat 5 butir soal yang memenuhi kriteria tingkat kesukaran, tidak ada soal yang sesuai dengan criteria daya pembeda, terdapat 16 soal yang termasuk pada criteria validitas cukup, sedangkan untuk criteria validitas tinggi hanya terdapat 1 soal. Hasil analisis instrument alat penilaian ujian kenaikan kelas ini termasuk dalam criteria reliabilitasnya rendah yaitu koefisien reliabilitasnya hanya 0,41.

Penelitian ini juga mengkaji instrument penilaian ujian kenaikan kelas yang dikaji dari segi kualitatif yang menunjukkan bahwa semua soal memenuhi criteria soal yang sesuai dengan indicator, criteria hanya ada satu jawaban yang tepat dan criteria kebenaran materi. Kajian perangkat kontruksi terdapat semua soal memenuhi criteria pokok soal tidak memberi petunjuk ke arah kunci jawaban, kriteria pokok soal bebas dari pernyataan yang bersifat negative ganda, criteria panjang rumusan relative sama, criteria pilihan jawaban tidak menggunakan pernyataan yang berbunyi semua jawaban diatas benar atau semua pilihan diatas salah, dan semua soal memenuhi criteria butir soal tidak tergantung pada jawaban sebelumnya. Sedangkan kajian dari perangkat bahasa terdapat tiga soal yang tidak memenuhi criteria soal menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia.

Relevansi penelitian ini dengan penelitian yang sedang dilakukan adalah sama-sama meneliti tentang instrument penilaian, tetapi terdapat perbedaan yaitu penelitian ini mengacu pada evaluasi alat penilaian sedangkan penelitian yang akan dilakukan adalah mengacu pada kualitas instrument penilaian. Dengan begitu dapat menjadi referensi yang tepat untuk peneliti.

### 1.3 KERANGKA KONSEPTUAL



Dalam kualitas instrument penilaian terdapat dua kategori yang dapat menjadi focus penelitian. Yang pertama meneliti secara kualitatif yang mencantumkan tentang bahasa, materi dan konstruksi yang terdapat dalam setiap butir soal. Yang kedua yaitu penelitian secara kuantitatif, dalam penelitian ini akan melakukan penelitian secara kuantitatif dan meneliti tingkat kesukaran, daya beda, dan pengecoh pada masing-masing soal.