

BAB II KAJIAN TEORI

A. Deskripsi Teoritis

1. Evaluasi Pendidikan

Evaluasi merupakan salah satu sarana penting dalam meraih tujuan belajar mengajar. Guru sebagai pengelola kegiatan belajar mengajar dapat mengetahui kemampuan yang dimiliki siswa, ketepatan metode mengajar yang digunakan, dan keberhasilan siswa dalam meraih tujuan pembelajaran yang telah ditetapkan melalui kegiatan evaluasi. Guru dapat mengambil keputusan secara tepat dengan informasi ini mengenai langkah apa yang harus dilakukan selanjutnya. Informasi tersebut juga dapat memberikan motivasi kepada siswa untuk berprestasi lebih baik.

Ujian *try out* pada hakikatnya merupakan evaluasi hasil belajar yang dilaksanakan oleh lembaga pendidikan sebelum menghadapi ujian nasional (UN). Hal-hal yang berkaitan dengan evaluasi belajar, seperti: pengertian, tujuan dan fungsi, prinsip evaluasi, dan jenis evaluasi akan diuraikan sebagai berikut.

a. Pengertian Evaluasi Pendidikan

Ada bermacam-macam pengertian evaluasi pendidikan yang dikemukakan oleh para ahli. Menurut Ratumanan (2003: 1), evaluasi dapat dinyatakan sebagai suatu proses sistematis dalam menentukan tingkat pencapaian tujuan instruksional. Ralph Tyler (dalam Arikunto, 2011: 3) mengatakan bahwa “Evaluasi merupakan sebuah proses pengumpulan data untuk menentukan sejauh mana, dalam hal apa, dan bagaimana tujuan pendidikan sudah tercapai. Jika

belum, bagaimana yang belum dan apa sebabnya”. Masih di dalam buku yang sama, definisi yang lebih luas dikemukakan oleh dua orang ahli, yakni Cronbach dan Stufflebeam. Tambahan definisi tersebut adalah bahwa proses evaluasi bukan sekedar mengukur sejauh mana tujuan tercapai, tetapi digunakan untuk membuat keputusan.

Jika evaluasi dikaitkan dengan pendidikan maka evaluasi pendidikan memiliki dua konsep pengertian. Hal ini sejalan dengan pendapat Sudijono (1996: 2) bahwa evaluasi pendidikan adalah: 1) Proses/kegiatan untuk menentukan kemajuan pendidikan, dibandingkan dengan tujuan yang telah ditentukan; 2) Usaha untuk memperoleh informasi berupa umpan balik (*feed back*) bagi penyempurnaan pendidikan.

Kesimpulan yang dapat diambil melalui beberapa konsep pengertian di atas, evaluasi pendidikan adalah suatu proses sistematis yang mengukur, menelaah, menafsirkan, dan mempertimbangkan sekaligus memberikan umpan balik (*feed back*) untuk mengetahui tingkat pencapaian terhadap tujuan pembelajaran yang telah ditetapkan serta digunakan sebagai informasi untuk membuat keputusan.

b. Tujuan dan Fungsi Evaluasi

1) Tujuan Evaluasi

Sudijono (1996: 16-17) menyatakan bahwa secara umum tujuan evaluasi belajar adalah untuk: (a) menghimpun bahan-bahan keterangan yang akan dijadikan sebagai bukti mengenai taraf perkembangan atau taraf kemajuan yang dialami oleh para peserta didik, setelah mereka mengikuti

proses pembelajaran dalam jangka waktu tertentu; dan (b) mengetahui tingkat efektivitas dari metode-metode pengajaran yang telah dipergunakan dalam proses pembelajaran selama jangka waktu tertentu.

Kegiatan evaluasi juga mempunyai tujuan khusus dalam bidang pendidikan, yaitu: (a) untuk merangsang kegiatan peserta didik dalam menempuh program pendidikan, dan (b) untuk menemukan faktor-faktor penyebab keberhasilan dan ketidakberhasilan peserta didik dalam mengikuti program pendidikan, sehingga dapat dicari dan ditemukan jalan keluar atau cara-cara perbaikannya.

2) Fungsi Evaluasi

Sudijono (1996: 7) menjelaskan bahwa secara umum ada tiga fungsi evaluasi, yaitu untuk: (a) mengukur kemajuan, (b) menunjang penyusunan rencana, dan (c) memperbaiki atau melakukan penyempurnaan kembali. Sudijono juga menambahkan, bahwa selain memiliki fungsi secara umum evaluasi juga memiliki fungsi secara khusus. Adapun fungsi evaluasi secara khusus dalam bidang pendidikan dapat ditinjau dari tiga segi, yaitu: (a) segi psikologi, (b) segi didaktik, dan (c) segi administratif. Lebih lanjut Sudijono menjelaskan tentang ketiga fungsi evaluasi tersebut, sebagai berikut.

Evaluasi pendidikan secara psikologi akan memberikan petunjuk untuk mengenal kemampuan dan status dirinya di antara kelompok atau kelasnya. Siswa akan mengetahui apakah dirinya termasuk berkemampuan tinggi, rata-rata, atau rendah. Apabila hal tersebut dapat dicapai maka

diharapkan evaluasi pendidikan akan dapat memberikan dorongan kepada siswa untuk memperbaiki, meningkatkan, dan mempertahankan prestasinya.

Evaluasi pendidikan bagi pendidik secara didaktik, setidaknya memiliki lima macam fungsi, yaitu: (1) memberikan landasan untuk menilai hasil usaha (prestasi) yang telah dicapai oleh peserta didik, (2) memberikan informasi yang sangat berguna untuk mengetahui posisi masing-masing siswa di antara kelompoknya, (3) memberikan bahan penting untuk memilih dan kemudian menetapkan status peserta didik, (4) memberikan pedoman untuk mencari dan menemukan jalan keluar bagi siswa yang memerlukannya, dan (5) memberikan petunjuk sejauh mana tujuan program pengajaran yang telah ditentukan telah dicapai.

Evaluasi pendidikan secara administrasi setidaknya memiliki tiga macam fungsi yaitu: (1) memberikan laporan mengenai kemajuan dan perkembangan siswa yang telah mengikuti kegiatan pendidikan dan pelatihan dalam jangka waktu tertentu, (2) memberikan bahan-bahan keterangan (data) untuk keperluan pengambilan keputusan, dan (3) memberikan gambaran mengenai hasil-hasil yang telah dicapai dalam proses pembelajaran.

c. Prinsip-Prinsip Evaluasi

Mengingat pentingnya evaluasi dalam menentukan kualitas pendidikan, maka upaya merencanakan dan melaksanakan evaluasi hendaknya memperhatikan beberapa prinsip. Menurut Daryanto (2005: 19-21), terdapat beberapa prinsip

yang perlu diperhatikan dalam melakukan evaluasi, yaitu keterpaduan, keterampilan siswa, koherensi, pedagogis, dan akuntabilitas.

1) Keterpaduan

Tujuan instruksional, materi, metode, pengajaran, serta evaluasi merupakan tiga kesatuan terpadu yang tidak boleh dipisahkan. Oleh karena itu, perencanaan evaluasi harus sudah ditetapkan pada waktu menyusun suatu pengajaran sehingga dapat disesuaikan secara harmonis dengan tujuan instruksional dan materi pengajaran yang hendak disajikan.

2) Keterlibatan Siswa

Untuk mengetahui sejauh mana siswa berhasil dalam kegiatan belajar mengajar yang dijalani secara aktif, siswa membutuhkan evaluasi. Penyajian evaluasi oleh guru merupakan upaya guru untuk memenuhi kebutuhan siswa akan informasi mengenai kemajuannya dalam program belajar mengajar. Siswa akan merasa kecewa apabila usahanya tidak dievaluasi.

3) Koherensi

Prinsip evaluasi dimaksudkan evaluasi harus berkaitan dengan materi pengajaran yang sudah disajikan dan sesuai dengan ranah kemampuan yang hendak diukur.

4) Pedagogis

Evaluasi dan hasil hendaknya dapat dipakai sebagai alat motivasi untuk siswa dalam kegiatan belajarnya.

5) Akuntabilitas

Evaluasi dan hasilnya dapat dipakai sebagai laporan pertanggungjawaban kepada pihak-pihak yang berkepentingan dengan pendidikan sehingga dapat diketahui sejauh mana keberhasilan pembelajaran yang telah dilakukan.

Dalam merencanakan dan melakukan evaluasi pembelajaran, seorang guru hendaknya selalu berpegang pada prinsip-prinsip tersebut. Hal ini dimaksudkan agar guru dapat bertindak dan berusaha seobjektif mungkin dalam mengadakan evaluasi.

d. Jenis Evaluasi

Menurut Daryanto (2005: 11-14) untuk masing-masing tindak lanjut yang dikehendaki dalam evaluasi diadakan tes yang disebut tes penempatan, tes formatif, tes diagnostik, dan tes sumatif.

Tes penempatan dilaksanakan pada awal tahun pelajaran baru, sebagai alat ukur untuk mengetahui tingkat kemampuan yang telah dimiliki peserta didik. Dengan demikian, siswa dapat ditempatkan pada kelompok yang sesuai dengan tingkat pengetahuan yang telah dimilikinya. Tes ini mengacu pada acuan norma.

Tes formatif dilaksanakan di tengah program pembelajaran untuk memantau kemajuan belajar siswa demi memberikan umpan balik, baik kepada siswa maupun kepada guru. Berdasarkan hasil tes tersebut dapat diketahui materi pelajaran apa yang belum dikuasai siswa sehingga guru harus mengupayakan perbaikannya. Tes ini mengacu pada acuan kriteria.

Tes diagnostik digunakan untuk mendiagnosa kesalahan belajar siswa dan mengupayakan perbaikannya. Pada jenis ini, tes formatif terlebih dahulu

disajikan untuk mengetahui ada tidaknya bagian mana yang belum dikuasai siswa, sehingga dapat dibuat butir-butir soal yang tingkat kesukarannya relatif rendah untuk mendekteksi.

Tes sumatif diberikan pada akhir tahun ajaran untuk memberikan nilai sebagai dasar menentukan kelulusan atau pemberian sertifikat bagi siswa yang telah menyelesaikan pelajaran dengan baik. Ruang lingkup tes sumatif mencakup seluruh bahan yang telah disajikan sepanjang jenjang pendidikan.

2. Instrumen Evaluasi

Dalam pengertian umum, alat adalah sesuatu yang dapat digunakan untuk mempermudah seseorang untuk melaksanakan tugas atau mencapai tujuan secara efektif dan efisien. Kata “alat” biasa juga disebut dengan istilah “instrumen”. Dengan demikian maka alat evaluasi juga dikenal dengan instrumen evaluasi (Arikunto, 2011: 25-26).

a. Bentuk Instrumen

Kegiatan evaluasi tidak dapat dipisahkan dari suatu instrumen evaluasi yang digunakan untuk mendapatkan data evaluasi. Untuk mendapatkan data evaluasi, evaluator menggunakan cara atau teknik. Cara tersebut dikenal dengan teknik evaluasi. Teknik evaluasi dapat dibedakan menjadi dua, yaitu teknik tes dan nontes. Kedua teknik tersebut digunakan untuk mencari data-data pengukuran tentang keberhasilan siswa dalam kegiatan belajar mengajar. instrumen yang banyak digunakan dalam bidang pendidikan ialah tes.

Tes adalah cara yang digunakan untuk melakukan pengukuran dalam bentuk tugas-tugas yang harus dikerjakan siswa agar mendapatkan data tentang nilai prestasi siswa tersebut untuk dibandingkan dengan nilai standar yang ditetapkan (Nurkencana dan Sumartana, 1983: 25). Bentuk tes dapat dibedakan menjadi dua, yaitu tes objektif dan tes subjektif.

Menurut Nurgiyantoro (2001: 11), tes subjektif digunakan agar siswa menunjukkan kemampuan dalam menerapkan pengetahuan, menganalisis, dan mengevaluasi informasi baru yang diharapkan kepadanya. Jawaban siswa terhadap tes subjektif menunjukkan kualitas cara berpikir siswa, aktifitas kognitif yang bukan semata-mata mengingat dan memahaminya saja.

Tes objektif adalah tes yang bersifat pasti, karena hanya ada satu kemungkinan jawaban yang benar. Sebagai alat pengukur hasil belajar siswa, tes ini sering dipilih karena dapat dikerjakan dengan cepat dan mencakup materi ajar yang luas, hanya memiliki satu jawaban yang pasti sehingga memudahkan korektor untuk mengoreksi, serta dapat dikoreksi secara cepat dengan hasil yang dapat dipercaya.

Menurut Arikunto (2011: 26) yang tergolong teknik nontes adalah skala bertingkat, kuesioner, daftar cocok, wawancara, pengamatan, dan riwayat hidup. Teknik nontes merupakan cara yang digunakan untuk memperoleh informasi yang berkaitan dengan tingkah laku kognitif. Nontes biasanya berisi pertanyaan atau pernyataan yang tidak memiliki jawaban benar atau salah secara mutlak.

b. Jenis Tagihan

Kegiatan evaluasi bentuk tes dan jenis tagihan yang digunakan hendaknya bervariasi tergantung pada jenis data yang ingin diperoleh. Dalam kaitannya dengan penelitian ini, jenis tagihan yang digunakan dalam ujian *try out* berupa soal tes pilihan ganda. Materi soal yang diujikan harus disusun sesuai dengan kisi-kisi soal yang telah dirumuskan.

3. Acuan Norma dan Kriteria

Acuan norma dan acuan kriteria menggunakan asumsi yang berbeda tentang kemampuan seseorang. Penafsiran hasil tes antara dua acuan ini berbeda hingga menghasilkan informasi yang berbeda maknanya.

a. Acuan Norma

Tes acuan norma berasumsi bahwa kemampuan setiap orang itu berbeda dan dapat digambarkan menurut distribusi normal. Perbedaan ini harus ditunjukkan oleh hasil pengukuran, misalnya setelah mengikuti tes peserta didik dibandingkan dengan kelompoknya.

Hal itu berarti standar yang diterapkan pada satu kelompok tidak dapat diterapkan pada kelompok lain. Standar yang dibuat berdasarkan hasil tes sebelumnya pun tidak dapat digunakan untuk hasil tes yang sekarang dan yang akan datang. Jadi, norma yang satu akan berbeda dengan norma yang lain. Kesemuanya bergantung pada prestasi siswa pada kelompok yang bersangkutan (Nurgiyantoro, 2001: 404).

b. Acuan Kriteria

Berbeda dengan pendekatan acuan norma, pendekatan acuan kriteria berusaha menafsirkan hasil tes yang ditetapkan standar yang akan dipergunakan sebelum kegiatan pengajaran dilakukan (Nurgiyantoro, 2001: 397).

Pada acuan kriteria seorang siswa dikatakan lulus apabila telah memenuhi tingkat penguasaan minimal pada bahan pelajaran. Konsekuensi dari acuan ini adalah adanya remidi dan pengayaan. Siswa yang belum memenuhi tingkat penguasaan minimal terhadap bahan pelajaran seperti yang diisyaratkan harus belajar lagi sampai kemampuannya mencapai standar yang telah ditetapkan. Sementara itu, bagi siswa yang telah mencapai standar diberi pelajaran tambahan yang disebut program pengayaan.

Soal ujian *try out* bersama mata pelajaran bahasa Indonesia kelas XII SMAN se-Kota Bima tahun pelajaran 2011/2012 disusun berdasarkan kisi-kisi soal yang mengacu pada Kurikulum Tingkat Satuan Pendidikan (KTSP). Sistem penilaipun mengacu pada kurikulum (KTSP). Sistem penilaian hasil pembelajaran KTSP menggunakan acuan kriteria dalam menafsirkan tes.

4. Validitas dan Reliabilitas Soal

Manfaat penilaian adalah untuk menentukan apakah sebuah tes telah mengukur apa yang hendak diukur dan apakah sebuah tes telah tepat digunakan untuk membuat suatu keputusan tentang pengambilan tes. Pada saat pengujian suatu soal, untuk mengambil keputusan tentang belajar siswa, harus dipastikan bahwa soal tersebut valid dan reliabel. Validitas soal dapat dilihat dari kisi-kisi

soal, sedang reliabilitas soal baru dapat diketahui setelah dilakukan uji coba terhadap soal tersebut. Uji coba ini dilakukan terhadap sejumlah subjek yang bersifat tipikal dengan populasi yang akan dites (Nurgiyantoro, 2001: 135).

a. Validitas Soal

Validitas atau (kesahihan) tes dapat diartikan sebagai ketepatan dan kecermatan tes dalam menjalankan fungsi pengukurannya. Suatu tes mempunyai validitas tinggi apabila tes tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dan tujuan diadakannya tes tersebut. Sebaliknya, suatu tes yang menghasilkan data tidak relevan dengan tujuan pengukuran dikatakan sebagai tes yang memiliki validitas rendah (Azwar, 1987: 146).

Menurut Arikunto (2011: 65-69) validitas sebuah tes dapat diketahui dari hasil pengalaman. Validitas tes tersebut dibagi menjadi dua kelompok, yaitu validitas logis dan validitas empiris. Validitas logis meliputi validitas isi (*content validity*) dan validitas konstruk (*construct validity*), sedangkan validitas empiris meliputi validitas “ada sekarang” atau konkruen (*concurrent validity*) dan validitas prediksi (*predictive validity*).

1) Validitas Isi (*content validity*)

Validitas isi didefinisikan sebagai kecocokan di antara isi alat ukur dengan sasaran ukur. Untuk keperluan pencocokan, biasanya isi sasaran ukur disusun dalam bentuk spesifikasi, yang meliputi bahan atau materi dan tujuan hasil belajar.

Sebuah tes dikatakan memiliki validitas isi apabila mengukur tujuan khusus tertentu yang sejajar dengan materi atau isi pelajaran yang diberikan. Oleh karena materi yang diajarkan tertera dalam kurikulum, maka validitas isi ini sering juga disebut dengan validitas kurikuler. Validitas isi ini diusahakan tercapai sejak saat penyusunan dengan cara merinci materi kurikulum atau materi buku pelajaran.

2) Validitas Konstruksi (*construct validity*)

Arikunto mengatakan bahwa sebuah tes memiliki validitas konstruk apabila butir-butir soal yang membangun tes tersebut mengukur setiap aspek berpikir seperti yang disebutkan dalam tujuan instruksional. Konstruk dalam pengertian ini bukanlah susunan, tetapi merupakan rekaan psikologis, yaitu suatu rekaan yang dibuat oleh para ahli ilmu jiwa yang dengan suatu cara tertentu merinci isi jiwa atas beberapa aspek seperti: ingatan, pemahaman, aplikasi, dan seterusnya.

3) Validitas Konkruen (*concurrent validity*)

Validitas ini lebih umum dikenal dengan validitas empiris. Sebuah tes dikatakan memiliki validitas empiris jika hasilnya sesuai dengan pengalaman. Dalam membandingkan hasil sebuah tes maka diperlukan suatu alat pembanding, maka hasil tes merupakan sesuatu yang dibandingkan.

4) Validitas Prediksi (*predictive validity*)

Menurut Arikunto sebuah tes dikatakan memiliki validitas prediksi atau validitas ramalan apabila mempunyai kemampuan untuk meramalkan apa yang terjadi pada masa yang akan datang.

Analisis kualitas soal *try out* bersama mata pelajaran bahasa Indonesia kelas XII SMA Negeri se-Kota Bima Tahun Pelajaran 2011/2012 menggunakan validitas isi. Validitas isi sering juga disebut validitas kurikulum, artinya bahwa suatu alat ukur dipandang valid apabila sesuai dengan kurikulum yang hendak diukur. Sebuah tes dapat dikatakan memiliki validitas isi apabila sesuai dengan tujuan khusus yang sama dengan isi pelajaran yang telah diberikan di kelas.

Untuk mengetahui apakah validitas isi telah dipenuhi atau belum, maka dilakukan telaah soal dengan cara menyesuaikan soal dengan kisi-kisi yang diacu. Dalam hal ini, digunakan *expert judgment* untuk menilai kesesuaian itu. Selain itu, untuk mengetahui validitas isi juga digunakan kartu telaah butir soal yang mencakup bidang kriteria penelaahan materi, konstruksi, dan bahasa.

b. Reliabilitas Soal

Reliabilitas suatu tes adalah konsistensi dari suatu tes dalam mengukur apa yang seharusnya diukur sehingga pengukuran itu memberikan informasi yang dapat dipercaya. Pengertian itu disimpulkan dari pendapat dua ahli pengukuran berikut ini. Arikunto (2011: 86) menyatakan bahwa reliabilitas dalam pengukuran berhubungan dengan masalah kepercayaan. Menurut Suryabrata (2004: 29), reliabilitas menunjukkan sejauh mana hasil pengukuran dengan alat tersebut dapat dipercaya. Hasil pengukuran harus reliabel artinya harus memiliki tingkat konsistensi dan kemampuan.

Suatu alat tes dikatakan reliabel apabila alat tes tersebut dapat dipercaya, konsisten, atau tetap. Untuk membuktikan apakah suatu alat tes memiliki sifat tetap, perlu diadakan uji coba terhadap alat tes yang akan digunakan tersebut.

Apabila sebuah tes diujicobakan lebih dari satu kali pada subjek yang sama dapat menghasilkan data yang kurang lebih sama, tes tersebut dikatakan dapat mengukur secara tetap.

Untuk mengukur reliabilitas soal dapat digunakan beberapa metode antara lain metode stabilitas, metode ekuavalensi, dan metode konsistensi internal (Surapranata, 2004: 90). Metode yang tepat digunakan pada penelitian ini adalah metode konsistensi internal. Metode konsistensi internal diperoleh dengan cara menganalisis data dari satu kali pengumpulan data (Widoyoko, 2009:147). Nunnally (dalam Surapranata, 2004: 113) menyatakan bahwa metode ini didasarkan pada korelasi antarskor jawaban pada setiap butir tes. Ia juga menyatakan bahwa metode ini khususnya digunakan pada setiap butir-butir yang dikotomi seperti pilihan ganda.

Reliabilitas soal dapat diketahui dari nilai *Alpha Cronbach* yang dihasilkan dari analisis kuantitatif *MicroCat Iteman*. Reliabilitas soal menggunakan metode konsistensi internal dapat dihitung dengan persamaan sebagai berikut:

$$r_{11} = \frac{n}{n-1} \left(1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right) \quad (1)$$

(Alpha via Arikunto, 2011: 109)

Keterangan:

r_{11} = reliabilitas yang dicari

n = jumlah soal

$\sum \sigma_i^2$ = jumlah varian skor tiap-tiap item

σ_t^2 = varian total

Menurut Nurgiyantoro (2004: 332) untuk tes-tes yang distandarkan, harga indeks reliabilitas itu paling tidak harus mencapai 0,85 sampai 0,90. Soal *try out* bersama mata pelajaran bahasa Indonesia ini disusun oleh tim MGMP (Musyawarah Guru Mata Pelajaran) sehingga koefisien reliabilitas yang relevan adalah sebesar 0,90 atau lebih dari 0,90. Semakin tinggi koefisien reliabilitas berarti semakin baik soal tersebut. Soal yang memiliki koefisien reliabilitas yang tinggi berarti kesalahan penyusunannya semakin kecil. Jumlah butir pertanyaan akan mempengaruhi indeks reliabilitas instrumen yang bersangkutan. Semakin banyak butir pertanyaan akan semakin tinggi indeks reliabilitasnya.

Reliabilitas soal dalam penelitian ini diuji dengan teknik reliabilitas *Alpha Cronbach*. Untuk membantu perhitungan digunakan komputer dengan program *MicroCat Iteman* versi 3.00.

5. Analisis Butir Soal

Kegiatan analisis butir soal dilakukan untuk mengetahui parameter soal yang terdiri dari indeks kesulitan dan keefektifan distraktor. Analisis butir soal dilakukan melalui program *MicroCat Iteman* versi 3.00.

a. Indeks Kesulitan

Arikunto (2011: 207) menyatakan bahwa “soal yang memiliki indeks kesulitan yang baik adalah soal yang tidak terlalu mudah tetapi tidak terlalu sukar”. Butir soal dikatakan mudah apabila sebagian besar siswa dapat menjawab dengan benar dan dikatakan sulit apabila sebagian besar siswa tidak dapat

menjawab dengan benar. Di dalam istilah evaluasi, indeks kesulitan ini diberi simbol P (p besar), singkatan dari kata “proporsi”. Besarnya indeks kesulitan (harga P) dapat dilihat dari proporsi peserta tes yang menjawab benar terhadap butir soal.

Tingkat kesulitan suatu butir soal dinyatakan dengan sebuah indeks yang berkisar antara 0,0 sampai dengan 1,0. Soal dengan indeks kesulitan 0,0 menunjukkan bahwa butir soal itu terlalu sulit, sebaliknya soal dengan indeks 0,1 menunjukkan bahwa soal yang bersangkutan terlalu mudah.

Penentuan indeks kesulitan (harga P) yang baik disesuaikan dengan jenis dan tujuan tes yang digunakan. Perbedaan jenis dan fungsi tes menghendaki harga P yang berbeda pula. Terkait dengan ketentuan tersebut, belum ditentukan harga p yang dianggap baik untuk suatu tes, perlu diketahui harga P tersebut akan dikenakan pada tes apa.

Indeks kesulitan yang sesuai untuk tes penempatan adalah sebesar 0,30-0,70; tes diagnosti dan tes sumatif sebesar 0,25 – 0,75 Gofur (dalam Khasanah, 2006: 28-29). Penelitian mengenai analisis soal *try out* ini tergolong tes sumatif sehingga indeks kesulitan yang digunakan sebesar 0,25-0,75. Pada analisis *Iteman*, indeks kesulitan butir soal ditunjukkan pada kolom *Proportion Correct*. Kriteria indeks kesulitan soal berkategori baik berada pada posisi 0,25-0,75, indeks kesulitan pada posisi $\leq 0,25$ berkategori sulit, dan indeks kesulitan $\geq 0,75$ tergolong dalam kategori mudah. Butir soal yang baik adalah butir soal yang memiliki indeks kesulitan sedang. Sedangkan butir soal yang memiliki indeks kesulitan mudah dan sulit, maka butir soal tersebut tergolong jelek sehingga perlu

direvisi atau diganti. Rumus untuk mencari indeks kesulitan adalah sebagai berikut.

$$P = \frac{B}{JS} \quad (2)$$

(Arikunto, 2011: 208)

Keterangan:

P = Indeks kesulitan

B = Banyaknya siswa yang menjawab soal dengan betul

JS = Jumlah seluruh siswa peserta tes

b. Indeks Daya Beda

Indeks daya beda mengkaji butir-butir soal dengan tujuan untuk mengetahui kesanggupan soal dalam membedakan antara siswa yang tergolong mampu dengan siswa yang tergolong kurang mampu (Nana, 2005: 141). Untuk mencari indeks daya beda butir soal, dapat menggunakan rumus sebagai berikut.

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B \quad (3)$$

(Arikunto, 2011:186)

Keterangan:

J = Jumlah peserta tes

JA = Banyaknya peserta kelompok atas

JB = Banyaknya peserta kelompok bawah

BA = Banyaknya kelompok atas yang menjawab benar

BB = Banyaknya kelompok bawah yang menjawab benar

Butir soal dinyatakan layak jika indeks daya bedanya baik. Indeks daya beda dapat berkisar antara -1,00 sampai dengan +1,00, tetapi jika indeks mendekati nol negatif maka dinyatakan tidak layak. Beberapa ahli mengatakan bahwa sebuah butir soal dikatakan layak apabila memiliki indeks daya beda sebesar 0,25 dan ada pula yang menyatakan 0,30. Penelitian ini melibatkan peserta ujian yang banyak, maka ditetapkan indeks daya beda yang layak minimum 0,30. Pada analisis iteman, indeks daya beda ditunjukkan oleh *Point Biser*.

Soal ujian *try out* bersama mata pelajaran bahasa Indonesia kelas XII tahun pelajaran 2011/2012 dimaksudkan untuk mempersiapkan siswa sebelum menghadapi ujian nasional (UN), itu artinya soal *try out* harus disusun berdasarkan kurikulum tingkat satuan pendidikan (KTSP). Acuan yang digunakan untuk menafsirkan hasil tes ini adalah acuan kriteria.

Indeks daya beda diabaikan karena acuan kriteria tidak menentukan perbedaan antara yang bodoh dan yang pintar. Dalam acuan kriteria pada prinsipnya semua siswa memiliki kemampuan yang sama dan bisa belajar apa saja, hanya waktu yang diperlukan untuk mencapai kemampuan tersebut berbeda-beda. Dalam Kurikulum Tingkat Satuan Pendidikan, indeks daya beda juga tidak digunakan, karena indeks daya beda akan dideskripsikan sebagai hasil penemuan analisis butir soal.

c. Keefektifan Distraktor

Tujuan analisis distraktor adalah untuk mengetahui seberapa banyak siswa menjawab benar sesuai dengan kunci jawaban dan seberapa banyak yang memilih distraktor atau pengecoh.

Menurut Fernandes via Suyata (dalam Nurwanti, 2008:28), suatu pengecoh dikatakan berfungsi dengan baik jika paling sedikit dipilih oleh 2% peserta tes. Dipihak lain, Azwar (2006: 151-152) menyatakan bahwa alternatif yang merupakan distraktor yang baik harus memiliki koefisien korelasi yang negatif, karena hal itu mengidentifikasikan bahwa pemilihannya memang berasal dari peserta didik yang kurang cakap. Dalam hal ini, analisis distraktor hanya dapat diterapkan pada tes pilihan ganda (tes objektif). Butir soal yang baik adalah butir soal yang kunci jawabannya positif, distraktornya negatif. Pada analisis *Iteman*, keefektifan distraktor ditunjukkan pada kolom *Proportion Endorsing*.

6. Karakteristik Pembelajaran Bahasa Indonesia di SMA Berdasarkan Kurikulum Tingkat Satuan Pendidikan

Kurikulum merupakan suatu komponen yang sangat penting dalam sistem pendidikan. Dalam Kurikulum Tingkat Satuan Pendidikan, secara garis besar penyusunannya meliputi pengembangan visi dan misi, perumusan tujuan satuan pendidikan, persiapan dan penyusunan draf, *review*, dan revisi. Kurikulum dijadikan acuan oleh setiap satuan pendidikan khususnya oleh guru dan kepala sekolah.

Dalam Mulyasa (2011: 4), kurikulum dibuat secara sentralistik, setiap satuan pendidikan harus melaksanakan dan mengimplementasikannya sesuai dengan petunjuk pelaksanaan dan petunjuk teknis yang disusun oleh pemerintah pusat yaitu Badan Standar Nasional Pendidikan (BSNP). Guru bertugas

menjabarkan kurikulum ke dalam satuan pelajaran sesuai dengan mata pelajaran masing-masing.

Pada Kurikulum Satuan Pendidikan, mata pelajaran bahasa Indonesia dikaitkan dengan standar kompetensi yang harus dicapai siswa. Standar kompetensi mata pelajaran bahasa Indonesia berorientasi pada hakikat pembelajaran bahasa dan sastra. Pembelajaran diarahkan untuk meningkatkan kemampuan siswa dalam berkomunikasi dengan Bahasa Indonesia, baik secara lisan maupun tertulis serta menimbulkan penghargaan terhadap hasil cipta manusia Indonesia.

Pembelajaran bahasa Indonesia di SMA menekankan pencapaian kemampuan berbahasa secara besar dan bukan pencapaian kemampuan tentang sistem bahasa. Di lain pihak, pembelajaran sastra lebih ditekankan pada pencapaian kemampuan apresiasi sastra daripada pengetahuan sastra sebagai ilmu. Hal tersebut juga berdampak pada sistem penilaian.

Standar kompetensi pelajaran bahasa Indonesia di SMA terdiri dari dua aspek kemampuan bersastra. Kedua aspek tersebut terbagi atas beberapa aspek mendengarkan, berbicara, membaca, dan menulis.

Pendekatan komunikatif dalam pembelajaran bahasa menekankan kemampuan berbahasa dan fungsi bahasa sebagai alat komunikasi pada pembelajaran tentang sistem bahasa. Pendekatan apresiatif dalam pembelajaran sastra lebih menekan pada apresiasi sastra yang dinyatakan pada kenyataan bahwa sastra merupakan salah satu bentuk seni yang bisa dinikmati atau diapresiasi.

Kompetensi yang ingin dicapai dalam pembelajaran bahasa Indonesia secara jelas telah termuat dalam Standar Kompetensi, yang kemudian dijabarkan ke dalam kemampuan dasar dan indikator. Berdasarkan indikator, kemudian disusun soal ujian yang digunakan untuk mengukur tingkat pencapaian kemampuan yang dimaksud.

B. Penelitian yang Relevan

Sumarsono (2000) melalui hasil analisis skripsi yang berjudul Analisis Kualitas Soal Pendalaman Materi Bahasa Indonesia SLTP di Kabupaten Sleman menyatakan bahwa validitas soal Pendalaman Materi Bahasa Indonesia SLTP di Kabupaten Sleman tergolong baik. Reliabilitas soal tersebut termasuk rendah (0.090). Dilihat dari indeks kesulitan soal secara keseluruhan termasuk rendah. Ditinjau dari indeks daya pembeda secara keseluruhan soal kurang baik. Analisis distraktor pada 70 soal pendalaman materi Bahasa Indonesia kelas III SLTP di Kabupaten Sleman Tahun Ajaran 1998/1999 secara keseluruhan tergolong baik.

Suryati (1998) melalui hasil analisis skripsi yang berjudul Karakteristik Ulangan Umum Bersama Bahasa dan Sasta Indonesia Kelas I Catur-Wulan II SMU Negeri se-Kabupaten Bantul Tahun Ajaran 1997/1998, menyatakan bahwa validitas isi soal ulangan umum bersama Bahasa dan Sasta Indonesia Kelas I Catur-Wulan II SMU Negeri se-Kabupaten Bantul Tahun Ajaran 1997/1998 tergolong baik. Reliabilitas soal tersebut termasuk rendah (0.408). Tingkat kesukarannya ditemukan bahwa 10 soal tergolong sulit (<0.25), 17 soal sedang ($0.25-0.75$), dan 18 soal mudah (>0.75). Jika dinyatakan dalam bentuk persentase,

soal sulit 22.2 %, soal sedang 37.8 %, dan soal mudah 4 %. Daya pembeda soal tersebut adalah 2 butir soal baik (>0.40), 15 butir soal sedang ($0.25-0.39$) dan 28 butir soal kurang (<0.25). Analisis distraktor pada 45 soal dengan alternatif sebanyak 225 ditemukan 170 baik sedangkan 55 kurang baik (51 direvisi dan 4 diganti).

Penelitian-penelitian tersebut menunjukkan bahwa soal sebagai alat ukur keberhasilan siswa belum disusun secara proporsional, masih perlu adanya evaluasi tentang penyusunan butir soal tersebut. Perbandingan semua aspek kognitif dalam soal juga masih perlu diperhatikan. Butir soal yang disajikan sebagai alat keberhasilan siswa pada tingkat yang lebih tinggi memuat kemampuan aspek kognitif yang tinggi pula. Penelitian-penelitian tersebut juga menunjukkan bahwa daya serap siswa terhadap pembelajaran bahasa Indonesia masih rendah sehingga perlu adanya peningkatan dalam segala aspek pembelajaran, baik pada proses pembelajaran maupun penilaian hasil belajar siswa.

C. Kerangka Pikir

Soal *try out* bersama mata pelajaran bahasa Indonesia kelas XII SMA Negeri se-Kota Bima Tahun Pelajaran 2011/2012 memiliki fungsi yang strategis terhadap prediksi hasil ujian nasional (UN). Oleh karena itu, sangat ideal apabila soal tersebut memiliki kualitas yang memenuhi persyaratan kaidah penulisan soal.

Dalam penelitian ini, soal tersebut dianalisis apakah sudah memenuhi persyaratan dari segi materi, konstruksi, bahasa, indeks kesulitan, dan fungsi

distraktor. Dari penelitian ini akan diperoleh soal yang sesuai dengan kaidah penulisan soal dan soal yang tidak sesuai. Soal yang memenuhi syarat kaidah yang ditentukan dapat dipakai sebagai contoh soal yang memenuhi syarat dan dapat dipakai sebagai alat evaluasi untuk ujian tahun berikutnya. Sedangkan soal yang tidak memenuhi syarat tentunya tidak layak dipakai sebagai alat evaluasi prestasi belajar atau soal tersebut harus direvisi atau bahkan diganti sama sekali.

D. Pertanyaan Penelitian

Penelitian ini bertujuan untuk mengetahui kualitas soal *try out* bersama mata pelajaran bahasa Indonesia kelas XII SMA Negeri se-Kota Bima Provinsi Nusa Tenggara Barat Tahun Pelajaran 2011/2012. Secara rinci dapat dirumuskan beberapa pertanyaan penelitian sebagai berikut.

1. Apakah butir soal *try out* bersama mata pelajaran bahasa Indonesia kelas XII SMA Negeri se-Kota Bima Tahun Pelajaran 2011/2012 ditinjau dari telaah butir soal telah memiliki validitas yang baik?
2. Bagaimanakah reliabilitas butir soal *try out* bersama mata pelajaran bahasa Indonesia kelas XII SMA Negeri se-Kota Bima Tahun Pelajaran 2011/2012 secara kuantitatif melalui proses *ITEMAN*?
3. Apakah analisis butir soal *try out* bersama mata pelajaran bahasa Indonesia kelas XII SMA Negeri se-Kota Bima Tahun Pelajaran 2011/2012 secara kuantitatif melalui proses *ITEMAN* telah memiliki indeks kesulitan dan keefektifan distraktor yang baik?